

2 802

BERICHTE
aus dem
INSTITUT FÜR MEERESKUNDE
an der
CHRISTIAN-ALBRECHTS-UNIVERSITÄT KIEL

Nr. 256

1994

DOI 10.3289/1FM-BER-256

Assimilation of Satellite Altimeter Data
into an Eddy-Resolving Primitive Equation Model
of the North Atlantic Ocean

von

Andreas Oschlies



Kopien dieser Arbeit können bezogen werden:
Institut für Meereskunde an der Universität Kiel
Abt. Theoretische Ozeanographie
Düsternbrooker Weg 20
24105 Kiel, -FRG-

ISSN 0341 - 8561 -

Diese Arbeit wurde von der Mathematisch-
Naturwissenschaftlichen Fakultät der
Universität Kiel als Dissertation angenommen.

Abstract

The aim of the present study is to develop an efficient assimilation scheme suitable for assimilating satellite altimeter data into a basin-scale eddy-resolving primitive equation model with active thermodynamics. Two alternative methods, referred to as extended nudging and re-initialization, respectively, will be presented and inter-compared. The first approach essentially extrapolates the surface observations into the ocean interior using *a priori* correlations provided by the model climatology. In contrast, the second assimilation procedure emphasizes dynamical relationships and in particular the conservation of temperature and salinity on isopycnals, rather than relying heavily on uncertain statistical correlations.

Both methods are tested and found to perform about equally well by running identical twin experiments, in which the model is assumed perfect and generates its own data for assimilation. Twin experiments are further used to compare the assimilation of the altimeter data directly along the satellite ground tracks with the assimilation of preprocessed maps combining all the data within some chosen time window. The results of these experiments indicate that – as far as the assimilation techniques proposed in this study are concerned – assimilation of quasi-synoptic maps is preferential.

The actual assimilation experiments are performed with objectively analyzed sea surface height anomalies measured by the U.S. Navy's satellite GEOSAT. Adding the model's mean sea surface to obtain an estimate of absolute dynamic sea level, these maps are then assimilated into the WOCE (World Ocean Circulation Experiment) CME (Community Modelling Effort) model of the North Atlantic Ocean at 5-day intervals covering the year 1987.

While the results of the two alternative assimilation procedures were almost indistinguishable in the identical twin experiments using simulated data, marked differences emerge when real altimeter data are used. These differences are shown to arise from different physical and statistical implications of the two assimilation methods. It turns out that a correct treatment of water mass properties is of crucial importance for the success of the assimilation scheme. Only the re-initialization method, which explicitly conserves temperature and salinity on isopycnals, leads to reasonable hydrographic situations, whereas the extended nudging scheme, that uses statistical correlations to infer temperature and salinity from sea surface height measurements, is found to introduce completely unrealistic water masses.

By comparing the results of both assimilation experiments with independent hydrographic observations, it is shown that the re-initialization method – in contrast to the extended nudging method – can reproduce the subsurface fields with an encouraging degree of accuracy. Not only does the assimilation of altimeter data significantly improve the model's representation of eddy activity, but it also leads to a more realistic description of major current systems like the North Atlantic Current and the Azores Current.

It is shown that combining surface restricted satellite data with a dynamical ocean circulation model can indeed provide new information about the state of today's ocean. The re-initialization method developed in this study is the first successful attempt to assimilate real altimeter data into a basin-scale eddy-resolving ocean circulation model with active thermodynamics. It is expected to be a useful tool for properly exploiting the wealth of forthcoming accurate altimeter data to monitor the changing ocean, thereby contributing to improve our understanding of its dynamics.

Contents

1	Introduction	1
2	The Use of Models in the Analysis of Altimeter data	3
2.1	Opportunities	3
2.2	Limitations	5
2.3	Methods	7
3	The Model	10
3.1	The CME Model	12
3.2	A simplified Box Configuration: High vertical Resolution	13
3.3	A simplified Box Configuration: Low vertical Resolution	15
4	Extended Nudging	18
4.1	The Nudging Approach	18
4.2	Extension to non-prognostic Variables	19
4.3	Intermittent Nudging	23
4.4	The Projection Operator \mathbf{H}	26
4.5	Nudging versus Optimal Interpolation	41
4.6	Choice of the Nudging Matrix \mathbf{M}	44
4.7	Summary	47

5	Re-Initialization	49
5.1	Beyond the Extended Nudging Approach	49
5.2	The Method	51
5.3	Static Performance Tests: A Snapshot	58
5.4	Dynamical Tests Using Identical Twin Experiments	64
6	The GEOSAT Data Set	68
6.1	Data Preprocessing	69
6.2	Maps versus Tracks	73
7	Results	81
7.1	Sea Surface Height	82
7.2	Eddy Kinetic Energy	92
7.3	Hydrography	99
7.4	Changes in the Model Climatology	108
8	Conclusions	113
A	Observability of Linear Systems	116
A.1	Application to Linear Models of the Ocean	119
	Bibliography	124

Chapter 1

Introduction

For most of the time throughout history of observational oceanography time variability in the ocean was mainly a nuisance, regarded to obstruct the determination of an unchanging ocean climate. Only during the last thirty years the importance of the low-frequency or mesoscale variability has been realized, favoured both by the use of advanced computers to run numerical ocean circulation models and by the development of instruments capable of obtaining long time series measurements at sea. Almost everywhere at mid and high latitudes ocean eddies can be observed, their kinetic energy generally exceeding that of the mean flow by more than an order of magnitude. They transport heat and water mass properties and interact with the general circulation in various ways not yet well understood. By now there is growing evidence that in the ocean there is no unchanging background climate that can be treated independently from the time-varying flow (MAROTZKE and WUNSCH, 1993). Intra- and interannual fluctuations as well as long-term changes in the ocean circulation are likely to have important effects on the global heat and hydrological cycles. Such phenomena have to be considered carefully in order to make reliable predictions about the Earth's changing climate.

To better understand and monitor the climatic role of time variability in the ocean, observations over long time spans and large areas are needed. While measurements carried out by ships and by instruments deployed in the sea provide essential oceanographic information, they are limited both in duration and in geographic coverage. Sensitive instruments mounted on Earth orbiting satellites can on the other hand take continuous measurements of almost any region of the world ocean. However, as sea water is opaque to electromagnetic radiation, such observations are inherently limited to the ocean surface. To obtain a comprehensive description of the complete ocean's

evolving state, one will therefore still have to exploit any information available, both by merging different data sets and also by employing reliable dynamical or statistical relationships between different variables.

One of the most important single data sets for studying global ocean dynamics is provided by satellite altimetry. By accurately measuring the variable part of the sea surface topography it is the only observing system that facilitates extensive sampling of mesoscale phenomena on a global basis. The suitability of satellite altimetry to describe the oceanic mesoscale variability has already been documented by a great number of studies (e.g., WILLEBRAND *et al.*, 1990; STAMMER, 1992; LETRAON, 1992). At present two remote sensing satellites (ERS-1 and TOPEX/POSEIDON) carry precise altimeters. Follow-on missions are scheduled to allow continuous monitoring of the world ocean throughout the near future.

In order to use this unprecedented wealth of accurate data correctly, it is important to fully appreciate potential benefits and limits of this data set. In particular, it has to be clarified to what extent these data can constrain sophisticated ocean circulation models. Such models, that represent much of our understanding of ocean dynamics, can efficiently be used to absorb the individual surface height measurements and dynamically interpolate the information in space and time. By using the dynamical constraints of the model one wishes to infer a complete state of the ocean's temperature, salinity and flow fields, that is consistent with the given time series of altimeter data.

The aim of this study is to establish a new assimilation scheme, specially adapted for assimilating satellite altimeter data into the most sophisticated ocean circulation models. It is probably true that only by coupling the forthcoming continuous flow of accurate data with such numerical models, oceanographers will be able to obtain a fully consistent description of the world ocean evolving in time. Eventually, the operational assimilation of any kinds of observations will provide an indispensable tool for monitoring, understanding and possibly predicting the changing climate of our ocean planet. The present work, which comprises the first successful assimilation of real altimeter data into a basin-scale eddy-resolving ocean model with active thermodynamics, is meant to be a basic contribution to this ambitious plan.

Chapter 2

The Use of Models in the Analysis of Altimeter Data



2.1 Opportunities

By its very conception a satellite altimeter can take observations of the ocean surface only. The instrument is a nadir-pointing pulse radar designed to make precise measurements of the time required for a pulse to travel from the altimeter to the surface and back. When combined with an accurate determination of the satellite's orbit as well as the atmospheric and ionospheric conditions influencing the velocity of the electromagnetic pulse, the topography of the ocean and also that of ice fields and flat lands can easily be mapped. The sea surface topography, however, consists of the geoid and the dynamic topography due to ocean currents. Since present geoids are not accurate enough to determine the absolute dynamic topography except at very long wavelengths of some thousand kilometers, oceanographers interested in mesoscale processes can only extract the time-varying part of the dynamic topography from the altimeter signal and use independent measurements to estimate the time mean.

A lot of work has been dedicated to the use of altimeter data for analyzing the mesoscale variability at the ocean surface (e.g., LETRAON *et al.*, 1990; STAMMER, 1992). Global statistical descriptions were obtained ranging from maps of the rms surface height variability to regional spectral analysis and characterization of typical space and time scales of mesoscale variability. These statistical descriptions have in common that they are essentially restricted to the ocean surface. It is only the use of our knowledge of ocean dynamics that may enable us to infer subsurface information from the altimetric data set.

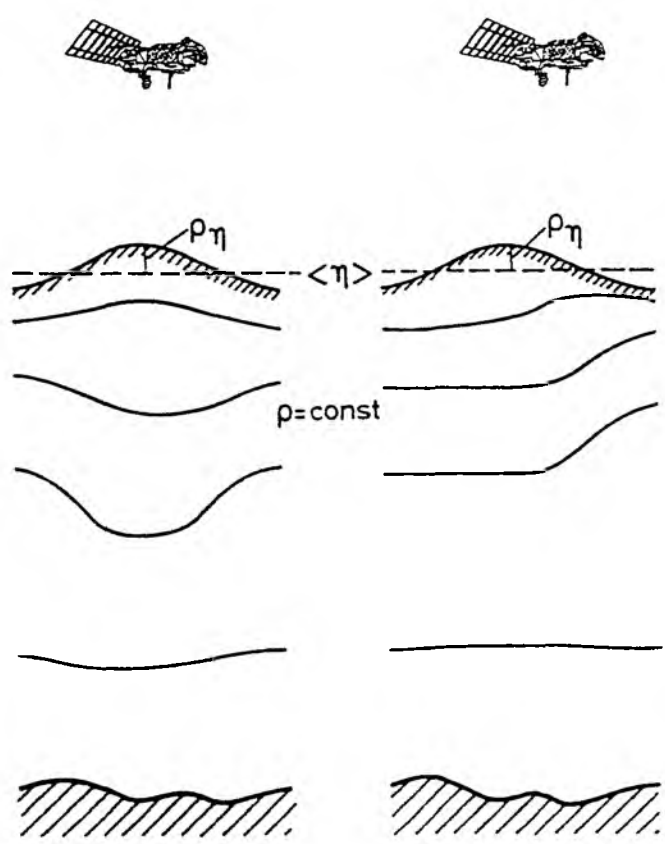


Figure 2.1: Symbolic representation of two water columns with identical surface topography but different density structure as indicated by isopycnals.

It is evident that a single observation of the sea surface topography alone will not allow a unique determination of the deep ocean's state. As shown in Figure 2.1 one can imagine very different states of the ocean producing the same instantaneous surface elevation. Though only two examples are depicted, in fact, there is an infinite number of such states indistinguishable with respect to the surface topography. However, due to the differences in the deep ocean structure, both states shown will evolve differently in time, eventually leading to distinct dynamic topographies. Hence, a second measurement taken by the satellite after some time will allow to restrict the number of possible states consistent with both observations. It therefore appears quite reasonable to expect that a long time series of surface observations combined with some understanding of ocean dynamics should gradually reveal more and more about the deep ocean. To what extent this expectation is legitimate shall be discussed in the following section.

Combining ocean dynamics and large amounts of data is usually best organized around a numerical circulation model including much of what today is known about the dynamical behaviour of the ocean. By assimilating the data into a numerical model, a huge number of observations can be dynamically extra- and interpolated both in space and in time. Particularly the model's ability to dynamically constrain the ocean interior given a time series of surface measurements is a very tempting prospect to globally monitor the state and, of course, also the potential changes of today's ocean. In this regard it seems appropriate to use one of the most realistic ocean circulation models available. Restricting ourselves to the study of the North Atlantic Ocean, we believe that at present the WOCE CME model, which is based on the Bryan-Cox code and will be described in chapter 3, is most appropriate for this objective. In the course of this work we shall therefore develop and run assimilation schemes suitable for this model.

Before starting any assimilation experiments, we shall nevertheless first look at the limits of this approach, what can be of some guidance for designing useful assimilation schemes and may also help to avoid misinterpretation of its results.

2.2 Limitations

It is considered appropriate to clarify to what extent one can expect benefits from assimilating altimeter data into numerical ocean circulation models. First of all, it is evident that the choice of the model is crucial. Clearly, models can only extract information about those dynamical modes supported by the model dynamics. For example, GASPAR and WUNSCH (1989), used a simple linear quasi-geostrophic model for the barotropic flow to assimilate six months of GEOSAT data of the North Atlantic. They showed that about 10% of the observed variability could be explained by only five barotropic Rossby waves. By increasing the number of possible wavemodes and moving to the more linear regime of the equatorial Pacific Ocean, FU *et al.* (1991) could after all explain 23% of the variance in the GEOSAT data set. One might wonder whether more realistic models would facilitate the extraction of even more information from the data, i.e., information that is possibly wasted by using too simple linear wave models.

However, when more complicated ocean circulation models like quasi-geostrophic layer models or even primitive-equation models with active thermodynamics are used in the assimilation process, it is not immediately obvious how well a time series of altimeter data can constrain the complete state of the three-dimensional flow and

density fields. The question arises whether there is additional gain of information that is worth the effort of using more and more refined models for assimilation studies.

This problem is usually investigated by running identical twin experiments, in which data generated by the model itself (by the so-called control run) are assimilated into the same model started from completely different initial conditions. If, during the assimilation, convergence towards the control run occurs, one generally assumes that the data fully constrain the model under consideration. Unfortunately, this test is not unambiguous, because it depends on the assimilation scheme chosen. For instance, BERRY and MARSHALL (1989) concluded that assimilating altimeter data into a three-layer quasi-geostrophic model was not sufficient to allow convergence towards the correct model state, in contrast to opposite results obtained by HOLLAND and MALANOTTE-RIZZOLI (1989) and also HAINES (1991), using a very similar model but different assimilation procedures. Having this in mind, one would like to have a more reliable criterion of the data set's ability to determine the state of the model.

Actually, such a criterion exists in form of the concept of observability. A dynamical system is said to be observable by a given observing scheme if there is a time span T such that the complete state at time t can be uniquely reconstructed from observations over the interval $[t - T, t]$. For linear systems a few general observability criteria have been established (e.g., COHN and DEE, 1988). A short excursion to the rather technical theory will be given in the Appendix, including some illustrative oceanographic examples and the derivation of an observability criterion best adapted for testing high-dimensional linear systems. Not unexpectedly, it turns out that pure inertia waves, which produce no surface height signal at all, are not observable by altimetry. It is further found that for linear layer models on the f -plane the degenerate geostrophic flow can not be observed by altimetry except for the very top layer, while on the β -plane the distinct evolution of different Rossby modes guarantees complete observability of the quasi-geostrophic flow, implying observability of the corresponding vertical displacements of isopycnals. However, when considering models with active thermodynamics it is obvious that changes in temperature and salinity which compensate with respect to density are not observable because they are not dynamically active. This "non-observability" of temperature and salinity individually will be explicitly accounted for by the re-initialization method developed in chapter 5.

Unfortunately, there is no well developed theory of observability for non-linear systems. It therefore is not immediately obvious how results obtained from linear theory translate to eddy-resolving circulation models, which are inherently non-linear. Only for those flow regimes which can be accurately described (over the time span T

introduced above, which for the present model would be of the order of years) by the linearized system, a theory of local observability can be derived. In general, non-linear effects will on the one hand transfer information between different modes, thus supporting observability, but on the other hand predictability may become a limiting factor. We shall conclude that, while examination of linear layer models indicates that quasi-geostrophic flow and density fields are generally observable by altimetry, data assimilation into non-linear models will still have to rely on convergence tests using identical twin experiments.

2.3 Methods

For the problem of assimilating observations into numerical ocean circulation models a variety of different methods already exists, most of them originally developed in meteorology. An extensive review can be found in GHIL and MALANOTTE-RIZZOLI (1991). While meteorological data assimilation techniques clearly provide some guidance to oceanographers, one has to keep in mind that there are major differences in the use of data in oceanography and meteorology. First, data coverage in meteorology is much more uniform and complete than in oceanography, with no observations at all for large regions of the deep ocean. Second, the motivation for data assimilation differs from the pressing need for accurate weather forecasts to the objective of better understanding the ocean circulation. Third, there are differences in the properties of the two fluids and also in the geometry and nature of the boundaries. Finally, numerical models of the atmosphere have reached a rather mature state, in contrast to present ocean circulation models which still have considerable shortcomings.

Taking into account the above mentioned points, we shall proceed by briefly discussing main assimilation techniques with respect to the objective of this study, i.e., developing a scheme suitable for assimilating satellite altimeter data into the eddy-resolving WOCE CME model of the North Atlantic Ocean.

a) Adjoint Method

Based on the calculus of variations, the adjoint method allows the determination of the one model trajectory that fits the sequence of observations best. Depending on the norm defined to measure the quality of the fit, one constructs a penalty or

cost function that has to be minimized with respect to some control variables, which mostly are the initial conditions of the model (LEDIMET and TALAGRAND, 1986). It is important to emphasize that adjusting only the initial conditions to find the best fit implicitly assumes that the model is perfect (LEWIS and DERBER, 1985), which may cause problems especially when applied to present ocean models. The minimization is efficiently carried out by essentially integrating the model forward and backward in time to compute gradients of the cost function. However, storage of the forcing terms as well as of non-linear effects is required over the full length of the time series of observations used. At least for basin-wide eddy-resolving circulation models this clearly goes beyond the storage capabilities of present computing systems. A further disadvantage of applying the adjoint method to non-linear systems is that increasing the assimilation period will rapidly increase the complexity of the shape of the cost function, making it virtually impossible to determine its global minimum. In this respect the introduction of more data will not always result in the “optimized” state being closer to reality (STENSRUD and BAO, 1992).

b) Kalman Filter

In contrast to the adjoint method the Kalman filter represents a sequential assimilation procedure (e.g., GHIL *et al.*, 1981). It is based on the statistical concept of optimal interpolation: At each observation time the Kalman filter optimally interpolates between the model forecast and the observations to obtain a new state vector with reduced error covariance. This state is subsequently used as initial state for the model to compute a forecast for the next observation time. By repeating this assimilation cycle and keeping track of the error covariance of the model state, the model step by step absorbs the information of the sequence of observations. The crucial point is that for this technique to be optimal one has to compute the time evolution of the covariance matrix of the model errors. It is again the corresponding computational burden which for our purpose clearly rules out the use of the Kalman filter.

c) Suboptimal Methods

Both methods mentioned above are statistically optimal in the sense that they minimize the covariance of the estimated errors. For linear models it can furthermore be shown that given the same information both the adjoint method and the Kalman

filter finally yield the same state estimate (THACKER, 1986). We note that the idea of minimizing the error covariance is not always adequate for present ocean models. If the probability distribution of the errors is not Gaussian – which in particular is the case for systematic model errors – the optimized state will in general not be the maximum likelihood one. In principle, one could even imagine cases in which suboptimal methods give state estimates which are more probable than those obtained from the so-called optimal methods. In any case, the huge computational cost immediately prohibits the application of these techniques to non-linear models with high-dimensional phase spaces. No simple statistically optimal method is available, and for our objective we therefore have to resort to suboptimal methods.

A large number of suboptimal methods not always equally well adapted to the assimilation problem (recall the opposite conclusions inferred from identical twin experiments by BERRY and MARSHALL (1989) and HOLLAND and MALANOTTE-RIZZOLI (1989) mentioned earlier) has been used in oceanography. In this work I will develop two alternative, suboptimal assimilation procedures, namely the extended nudging method described in detail in the fourth chapter, and a re-initialization approach presented in chapter 5. Their different statistical and dynamical implications shall be discussed with respect to the objective of improving our understanding of the ocean in general and the climatic role of mesoscale variability in particular.

Chapter 3

The Model

All the model calculations carried out in this study are based on the GFDL (Geophysical Fluid Dynamics Laboratory) numerical ocean circulation model developed by BRYAN (1969) and adapted for vector computers by COX (1984). The model essentially integrates the primitive equations (e.g., MÜLLER and WILLEBRAND, 1989), which are derived from the conservation equations for momentum, mass, heat and salt. Major simplifications introduced are the rigid-lid approximation, that filters out fast external gravity waves, the neglect of tidal effects, and the Boussinesq as well as the hydrostatic approximations. The latter implies that convection processes removing static instabilities have to be parameterized. In the model versions used, a convective adjustment process implemented by COX (1984) is applied that locally switches to a high vertical diffusivity ($10^4 \text{ cm}^2 \text{ s}^{-1}$).

The continuous equations of momentum, continuity, potential temperature T and salinity S can be written as

$$\frac{\partial \mathbf{u}_h}{\partial t} + \mathcal{L} \mathbf{u}_h + w \frac{\partial \mathbf{u}_h}{\partial z} + \mathbf{f} \times \mathbf{u}_h = -\frac{1}{\rho_0} \nabla p + K_M \frac{\partial^2 \mathbf{u}_h}{\partial z^2} + \mathbf{F}_M \quad (3.1)$$

$$\nabla \cdot \mathbf{u}_h + \frac{\partial w}{\partial z} = 0 \quad (3.2)$$

$$\frac{\partial p}{\partial z} = -\rho g \quad (3.3)$$

$$\frac{\partial T}{\partial t} + \mathcal{L} T + w \frac{\partial T}{\partial z} = K_D \frac{\partial^2 T}{\partial z^2} + F_T \quad (3.4)$$

$$\frac{\partial S}{\partial t} + \mathcal{L} S + w \frac{\partial S}{\partial z} = K_D \frac{\partial^2 S}{\partial z^2} + F_S \quad (3.5)$$

$$\rho = \rho(S, T, p) \quad (3.6)$$

where \mathbf{u}_h is the vector of horizontal velocity (u, v) , w vertical velocity, and \mathcal{L} refers to the horizontal advection operator. The vector \mathbf{f} is the Coriolis parameter times a unit vector in local z -direction, ρ is the density computed either by the polynomial approximation of BRYAN and COX (1972) (sections 3.1 and 3.3) or a linear state equation neglecting salinity (section 3.2), g is the acceleration due to gravity, and p is pressure. K_M and K_D are the vertical mixing coefficients for momentum and tracers, respectively, and \mathbf{F}_M and $\mathbf{F}_{(T,S)}$ parameterize the effects of lateral turbulent mixing.

At lateral walls a “no slip” ($\mathbf{u} = 0$) boundary condition is imposed, while at the bottom a quadratic bottom drag acts with a drag coefficient of $c_D = 1.3 \times 10^{-3}$. There is neither heat nor salt flux through the bottom or lateral walls. Fluxes of heat and salt across the ocean surface are generally realized in form of restoring conditions. By setting $w = 0$ at $z = 0$, the rigid-lid approximation eliminates external gravity waves and hence allows for larger model timesteps. Instead of surface elevation (or surface pressure) the model therefore uses the volume transport streamfunction Ψ as prognostic variable: By vertically integrating the momentum equation (3.1) and subsequently taking the curl, the surface pressure gradient is eliminated and a prognostic equation for $\nabla \cdot (H^{-1} \nabla \Psi)$ can be obtained, which is solved for Ψ by a Poisson solver, using the kinematic boundary condition $\Psi = \text{const}$, and the “no slip” boundary condition $\partial \Psi / \partial n = 0$.

Of particular importance for the assimilation of satellite altimeter data is, of course, the surface pressure field, which, as just pointed out, is not a prognostic quantity of the model. The surface pressure gradient can, however, readily be recovered diagnostically from prognostic model variables (e.g., SEMTNER, 1986). The surface pressure field itself may then be obtained by integrating its gradient over the model domain. The remaining integration constant is the mean sea surface elevation, which for the reason of volume conservation is set to zero. Note that as a particular result of this integration procedure, the diagnosed surface height depends on the model state variables in a highly non-local way. It is essentially this non-locality of the SSH field that severely obstructs a straightforward assimilation of altimeter data into sophisticated ocean circulation models.

3.1 The CME Model of the North Atlantic Ocean

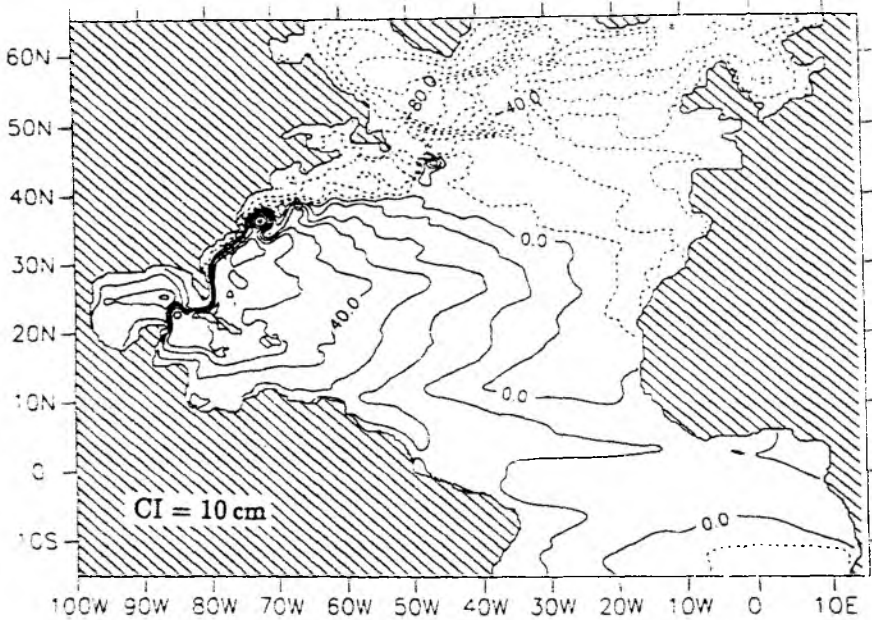


Figure 3.1: Mean sea surface height of the CME model configuration (averaged over three years). The corresponding rms variability is shown in Figure 7.2.

The recent decades have shown that it is indeed possible to simulate the ocean circulation with some degree of realism. A number of studies have already been aimed at direct comparisons of observations of the real ocean with the output from numerical models. One of the models that so far has been used quite successfully (SPALL, 1990; SCHOTT and BÖNING, 1991; BÖNING *et al.*, 1991; DIDDEN and SCHOTT, 1992) is the numerical model of the North and equatorial Atlantic Ocean set up within the Community Modelling Effort (CME) of the World Ocean Circulation Experiment (WOCE). It was explicitly designed for investigating the role of eddies in the circulation of realistic ocean basins, driven by realistic forcing fields (BRYAN and HOLLAND, 1989). The model is based on the same primitive-equation code described above. The grid spacing is $1/3^\circ$ in latitude and $2/5^\circ$ in longitude, and 30 levels in the vertical with a spacing increasing from 35 m at the surface to 250 m near the bottom. The model domain covers the Atlantic Ocean between 15°S and 65°N (Figure 3.1) with realistic bottom topography. Northern and southern lateral boundaries are closed as is the Strait of Gibraltar, with buffer zones (“sponge layers”) with a width of five grid points at which the temperature and salinity of the water is relaxed to values, that in the present

experiment were taken from the LEVITUS (1982) atlas.

The horizontal subgrid-scale dissipation and viscosity are represented by biharmonic operators, $\mathbf{F}_M = -A_M \nabla^4 \mathbf{u}_h$, and $F_{(T,S)} = -A_D \nabla^4 (T, S)$, with coefficients $A_M = A_D = -2.5 \times 10^{19} \text{ cm}^4 \text{ s}^{-1}$. Vertical viscosity is represented by the Laplacian operator with a coefficient of $K_M = 10 \text{ cm}^2 \text{ s}^{-1}$, while for the vertical diffusion of tracers a coefficient of $K_D = 0.3 \text{ cm}^2 \text{ s}^{-1}$ was chosen. A simple mixed layer model (after CAMP and ELSEBERRY, 1978) computes the depth of vertically uniform temperature and salinity from the vertical flux of kinetic energy in the atmospheric boundary layer.

The model was originally started from rest with initial temperature and salinity fields taken from the Levitus climatology. The experiment used in this study is forced with monthly mean wind stresses of ISEMER and HASSE (1987) north, and of HELLERMAN and ROSENSTEIN (1983) south of the equator (with a transition zone between 0° and 5°N). The thermohaline forcing is represented by a relaxation of surface salinity to the monthly mean values of LEVITUS (1982) and a heat flux given by the linear formulation of HAN (1984). For the experiments to be described in this study, the model years 30 to 32 were used as climatology. The actual assimilation of GEOSAT data of the year 1987 was performed in model year 32.

3.2 A simplified Box Configuration: High vertical Resolution

Because of the large computational requirements of running the CME model of the North Atlantic Ocean – 18 MW memory and 40 CPU hours per model year on a CRAY-XMP – it was decided to use somewhat simplified models for all test calculations during the development of the assimilation schemes. The first model being used is the box model configuration of COX (1985) with rough bottom topography as described by BÖNING (1989). It is based on essentially the same numerical code as the CME model, the only change in the physics being a simplified state equation: There is no salt in the model, and density ρ is a linear function of the single state variable potential temperature T only. The model is set up for a simplified basin of 60° longitudinal width, extending from the equator to 65°N (Figure 3.2 *b*). There are 21 levels in the vertical and the horizontal resolution is $1/3^\circ \times 2/5^\circ$. Besides the reduced number of model grid points it is also the neglect of salinity that reduces the computational requirements to about 6 MW and less than 8 hours of CPU time per model year on a CRAY-XMP.

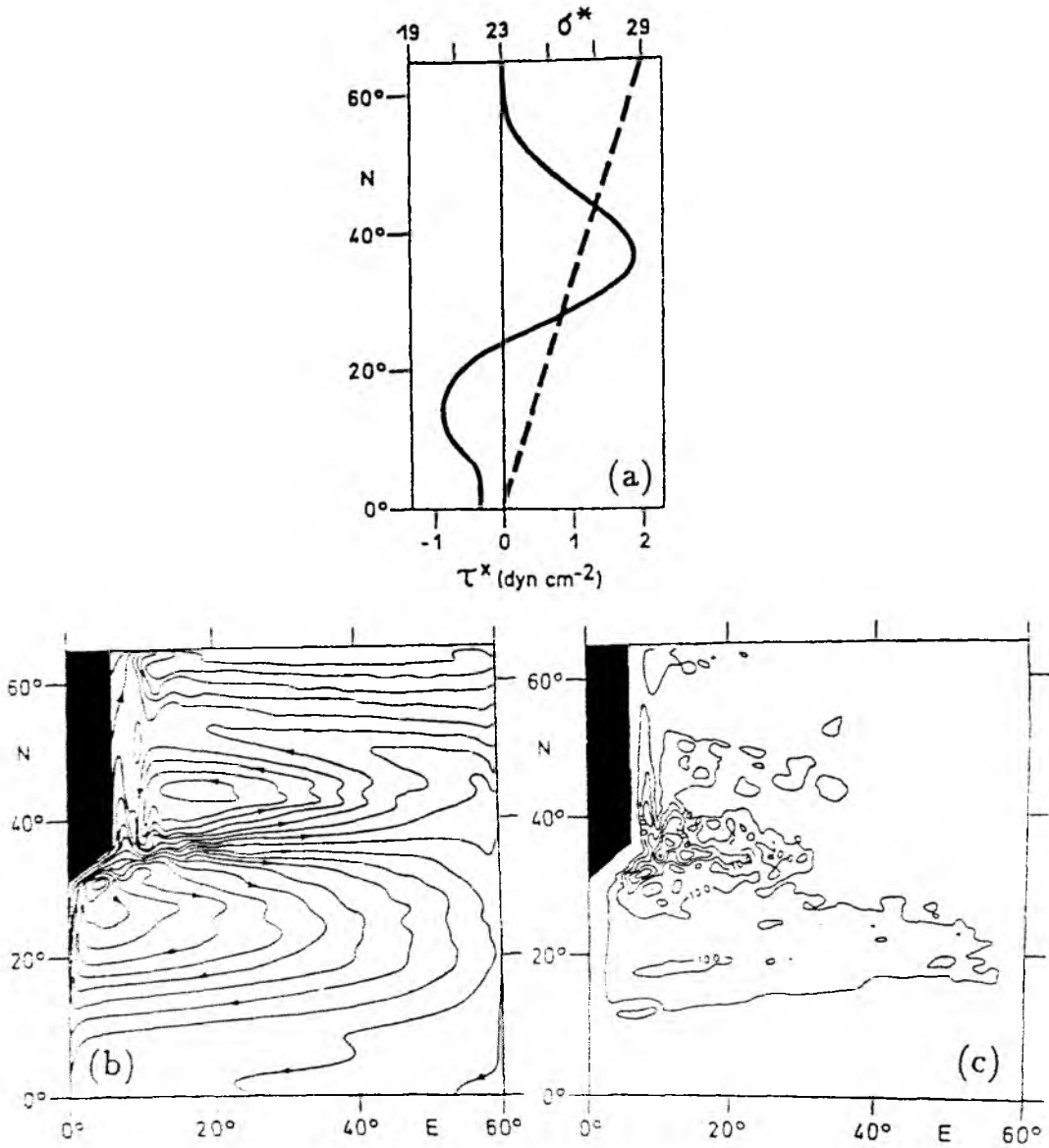


Figure 3.2: The idealized box model configuration of COX (1985): (a) zonal wind stress (solid line) and reference density (dashed line) at the surface, (b) 5-year mean surface elevation with contour interval = 10 cm (from BÖNING and BUDICH, 1992), and (c) surface variability with contour interval = 5 cm rms.

The vertical mixing coefficients chosen are identical to those of the CME model, while the horizontal mixing coefficients are slightly changed to $A_M = -0.8 \times 10^{19} \text{ cm}^4 \text{ s}^{-1}$, and $A_D = -2.4 \times 10^{19} \text{ cm}^4 \text{ s}^{-1}$. The box ocean is forced by a steady zonal “double gyre” wind stress τ and a surface density relaxation to a constant meridional gradient (Figure 3.2 *a*) with a timescale of 50 days.

The model years 30 to 32 are taken as climatology. Complete fields of the diagnostically calculated surface pressure as well as the corresponding three-dimensional fields of velocity components and temperature are stored twice a day for the model year 30. In the identical twin experiments these data from the so-called control run serve as substitute for the real world. The surface pressure maps are subsequently assimilated into the assimilation run that starts from different initial conditions at the beginning of model year 31. An objective criterion for the convergence of the assimilation process can be obtained by computing the rms difference between the fields of the assimilation run and those stored from the control run. Because the simulated observations are dynamically and statistically consistent with the assimilation model, the results of such twin experiments tend to be over-optimistic with respect to applications to real data. However, because the history of the control run, i.e., the substitute for the real world, is completely known, identical twin experiments provide an ideal tool for developing and testing assimilation schemes.

3.3 A simplified Box Configuration: Low vertical Resolution

Compared with the state space dimension of the CME model of about 6.4×10^6 (four prognostic variables at each of the 1.6×10^6 grid points), the 21-layer box model version of the previous section with a state dimension of about 1.9×10^6 already significantly reduces the computational costs of running the model. However, in order to allow for a greater number of identical twin experiments a further simplified box model version was set up. With merely four levels in the vertical (with box thickness of 100 m, 250 m, 750 m, and 2500 m, respectively) and a basin extending from 20°N to 44°N and a longitudinal width of 27° (Fig. 3.3 *b*), the state dimension of this $2/5^\circ \times 1/3^\circ$ eddy resolving model amounts to 5.8×10^4 only. In contrast to the previous models, here the grid spacing is $1/3^\circ$ in zonal and $2/5^\circ$ in meridional direction. Without having much effect on the circulation, this grid simplifies the generation of a synthetic GEOSAT orbit (see chapter 6). A flat bottom is chosen at 3600 m depth. For the

coarse vertical resolution runs to become not too unstable, the model's viscosity had to be increased. The mixing parameters finally used are $A_M = A_D = -5 \times 10^{19} \text{ cm}^4 \text{ s}^{-1}$, $K_M = 1000 \text{ cm}^2 \text{ s}^{-1}$ and $K_D = 1.0 \text{ cm}^2 \text{ s}^{-1}$.

Both the “double gyre” wind stress and the restoring surface density (50 days relaxation timescale) are functions of latitude only and are shown in Figure 3.3 *a*. The model is started from horizontally homogeneous temperature and salinity profiles that reflect conditions in the subtropical Atlantic (18°C, 35.5 psu; 15°C, 35.3 psu; 9°C, 35.1 psu; 4°C, 34.9 psu). Since the reduced model domain does not allow for a formation of realistic deep water, the bottom layer is restored to the initial T-S values (4°C, 34.9 psu) on a timescale of 100 days. Again, model years 30 to 32 provide the climatology. For the identical twin experiments year 30 serves as control run, while the assimilation runs start at the beginning of year 32. This model version will be used mainly for the space-versus-time resolution studies of chapter 6, comparing along-track assimilation with assimilation of preprocessed maps.

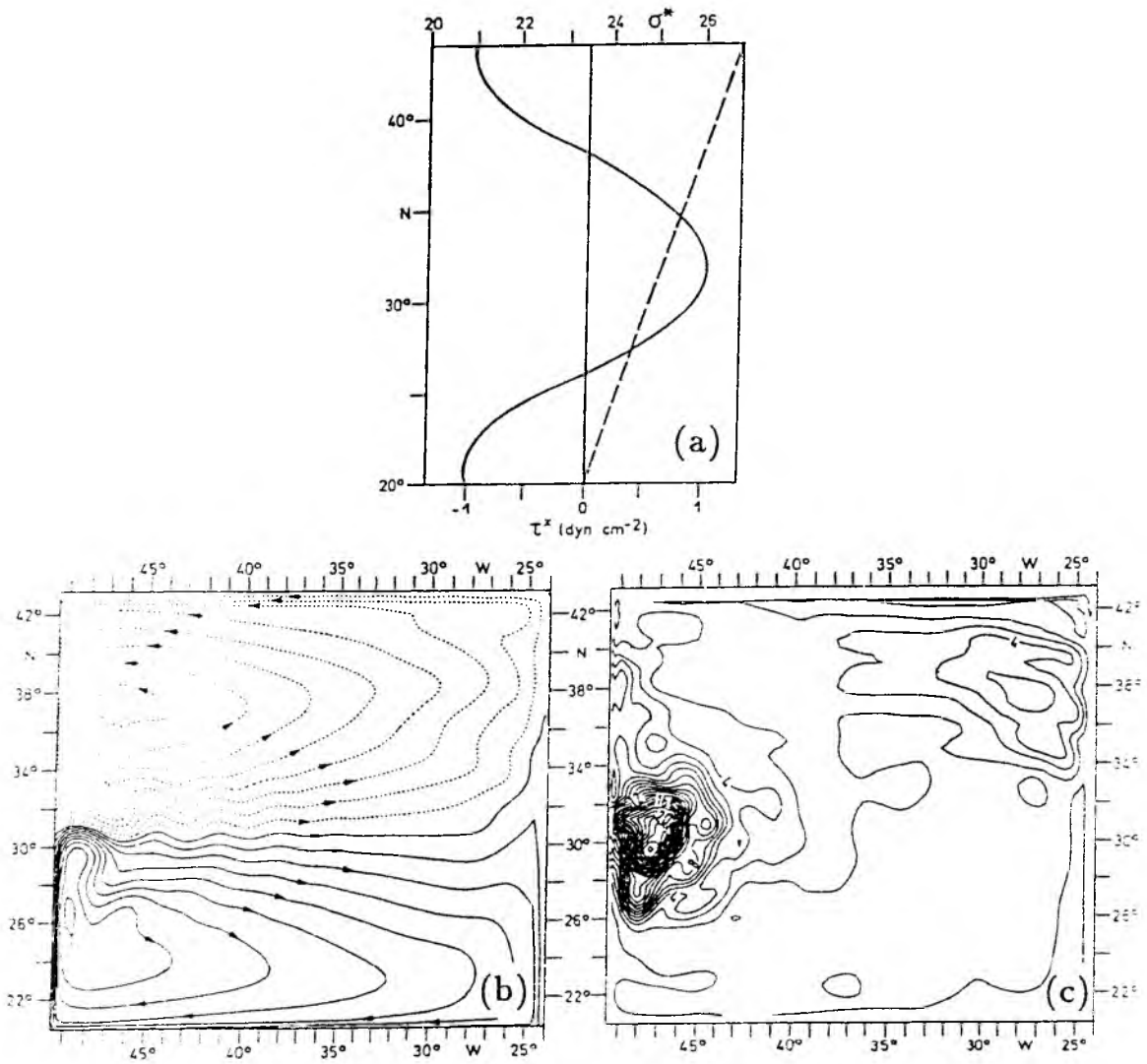


Figure 3.3: The box model configuration of the 4-level model: (a) zonal wind stress (solid line) and reference density (dashed line) at the surface, (b) 3-year mean surface elevation with contour interval = 6 cm, and (c) surface variability with contour interval = 1 cm rms.

Chapter 4

Extended Nudging

4.1 The Nudging Approach

Despite many theoretical investigations aimed towards the development of so-called optimal assimilation methods, an operational application of these advanced techniques to a full-scale general circulation model requires computational resources not available today. For this reason, here we shall first of all consider the computationally and conceptually simple, but statistically suboptimal nudging technique which, however, has a rather poor theoretical foundation. It is my objective to clarify some of the attributes of this technique and to establish an extended nudging scheme that as far as possible avoids any empirical tuning, involved in most previous applications of the nudging technique.

First employed by ANTHES (1974) to dynamically initialize hurricane prediction models, the nudging or Newtonian relaxation technique since then has been widely used in meteorology and also oceanography. The basic concept of this assimilation method is given for a scalar, directly observed state variable, x , by

$$\frac{\partial x}{\partial t} = \text{“physics”} + \mu(x^{obs} - x) \quad , \quad (4.1)$$

where in addition to the model physics a linear relaxation term is introduced which nudges the prognostic model variable towards the observed value x^{obs} . The positive, inverse relaxation timescale μ , also called nudging parameter, has to be chosen properly since there is a subtle balance between forcing the model towards the observations and disturbing the model physics by the non-physical nudging term $\mu(x^{obs} - x)$. A useful criterion how to best choose the nudging parameter will be developed later in this chapter.

A number of oceanographic applications of the nudging technique has been concentrated on the assimilation of altimeter data into quasi-geostrophic models (e.g., HOLLAND, 1989; HOLLAND and MALANOTTE-RIZZOLI, 1989; HAINES, 1991; VERRON, 1992). These models have in common that the first layer vorticity, which via geostrophy directly relates to the sea surface height as observed by altimetry, is a prognostic variable of the model. Because the primitive-equation CME model we will use has a rigid lid to exclude fast external gravity waves, sea surface height is not a prognostic variable, but can be calculated diagnostically only (see chapter 3). Hence a first major problem to be faced is to extend the standard nudging technique to cases where the model counterpart of the observations is not given simply by some prognostic model variables, but instead is a more complicated function of the state vector.

4.2 Extension to non-prognostic Variables

A rather general way – only assuming a linear relationship between observations and state variables – to extend the scalar nudging approach (4.1) to the assimilation of any vector of observations, \mathbf{y}^{obs} , into a physical system with state vector \mathbf{x} can be formulated as

$$\frac{\partial \mathbf{x}}{\partial t} = \text{“physics”} + \mathbf{N} \mathbf{y}^{obs} - \mathbf{M} \mathbf{x} , \quad (4.2)$$

including the case that observations do not directly correspond to prognostic model variables. It is not immediately obvious how to choose the two matrices \mathbf{N} and \mathbf{M} . In general, both matrices can vary with time, depending on the observation schedule. For the moment we shall simplify the analysis by assuming continuous observations of always the same kind. This allows us to consider constant matrices \mathbf{N} and \mathbf{M} only, and we defer the problem of intermittent observations to the following section 4.3.

Useful information about the structure of these matrices, necessary for the concept of assimilation to work, can be obtained by analyzing the effect the nudging procedure has on linear systems like

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{A} \mathbf{x} + \text{“forcing”} + \mathbf{N} \mathbf{y}^{obs} - \mathbf{M} \mathbf{x} , \quad (4.3)$$

where the forcing is assumed to be known, and the nudging term $\mathbf{N} \mathbf{y}^{obs} - \mathbf{M} \mathbf{x}$ has already been included. For perfect observations combined with perfect model physics it is adequate to require

$$\mathbf{N} \langle \mathbf{y}^{obs} \rangle = \mathbf{M} \mathbf{x}^{true} , \quad \frac{\partial \mathbf{x}^{true}}{\partial t} = \mathbf{A} \mathbf{x}^{true} + \text{“forcing”} , \quad (4.4)$$

where the expectation value $\langle y^{obs} \rangle$ has been used to filter out observational noise. According to the superposition principle we can subtract this particular solution from the actual trajectory of the model state and obtain an equation for the evolution of the expected model error $\mathbf{x}' = \langle \mathbf{x} - \mathbf{x}^{true} \rangle$ only:

$$\frac{\partial \mathbf{x}'}{\partial t} = \mathbf{A} \mathbf{x}' - \mathbf{M} \mathbf{x}' . \quad (4.5)$$

A necessary condition for convergence of the assimilation procedure towards truth is then $\lim_{t \rightarrow \infty} \mathbf{x}'(t) = 0$, for any initial $\mathbf{x}'(0)$. This is equivalent to the condition that all the eigenvalues of the new system matrix $(\mathbf{A} - \mathbf{M})$, i.e., the original system matrix \mathbf{A} disturbed by the action of the nudging matrix \mathbf{M} , have negative real parts.

Quite often, linear ocean models simply represent some superposition of wave solutions, and we shall therefore assume the system matrix \mathbf{A} be diagonalizable with purely imaginary eigenvalues. Then a transformation matrix \mathbf{S} , its columns being the eigenvectors of \mathbf{A} , exists, with $\mathbf{A} = \mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix containing the imaginary eigenvalues. By inserting $\mathbf{S}^{-1} \mathbf{x}'$ and its complex conjugate $(\mathbf{S}^{-1} \mathbf{x}')^\dagger$ into equation (4.5) it is then straightforward to derive

$$\frac{\partial}{\partial t} |\mathbf{S}^{-1} \mathbf{x}'|^2 = - \left(\mathbf{S}^{-1} \mathbf{x}' \right)^\dagger \left(\left(\mathbf{S}^{-1} \mathbf{M} \mathbf{S} \right) + \left(\mathbf{S}^{-1} \mathbf{M} \mathbf{S} \right)^\dagger \right) \left(\mathbf{S}^{-1} \mathbf{x}' \right) . \quad (4.6)$$

From this result we conclude that, in order for the energy of the linear-wave model misfit to decrease during assimilation, the symmetric and real matrix $\left(\left(\mathbf{S}^{-1} \mathbf{M} \mathbf{S} \right) + \left(\mathbf{S}^{-1} \mathbf{M} \mathbf{S} \right)^\dagger \right)$ has to be a positive definite operator. Because for high-dimensional systems (if they are diagonalizable anyway) both computation of the transformation matrix \mathbf{S} as well as demands on storing a full matrix \mathbf{M} will be rather forbidding, to be on the safe side, one will usually choose a positive diagonal matrix \mathbf{M} and nudge the respective prognostic variable only (but see also VERRON, 1992).

To illustrate the importance of the nudging matrix \mathbf{M} having a proper structure, I will give a noteworthy though simple oceanographic example. Having in mind that later in this study we shall use the CME model which, due to the rigid-lid approximation, has no prognostic equation for the surface elevation, one might try to assimilate satellite altimeter data by transferring the nudging term to the momentum equations. This obviously requires the use of some diagnostic relationship between the observed variable sea surface height, η^{obs} , and the prognostic velocity components, u and v , of the model. Because the spatial and temporal sampling characteristics of satellite altimeters are generally best suited for observations of mesoscale processes (see chapter 6), we restrict ourselves to the geostrophic part of the observations and neglect

any signal due to inertia-gravity waves. The assumption of geostrophic balance is then used to directly relate observations of the sea surface height to “observed” velocity components. Still, two fundamentally different approaches of constructing the nudging terms – and hence \mathbf{M} – exist:

First, it seems natural to use the same geostrophic relation, that transforms the actual measurements to “pseudo” observations of current velocities, for the model variables as well and to write

$$\frac{\partial u}{\partial t} = \text{“physics”} + \mu \left(-\frac{g}{f} \frac{\partial}{\partial y} (\eta^{obs} - \eta) \right) \quad (4.7)$$

$$\frac{\partial v}{\partial t} = \text{“physics”} + \mu \left(\frac{g}{f} \frac{\partial}{\partial x} (\eta^{obs} - \eta) \right) , \quad (4.8)$$

with some positive nudging parameter μ . This approach exclusively compares the geostrophic part of the model’s velocity field with the observations. Due to the appearance of η in the nudging terms of the momentum equations, this approach corresponds to a non-diagonal nudging matrix \mathbf{M} . Indeed, test calculations with the GFDL model using the above nudging approach quickly led to a sharp increase in the total kinetic energy that finally destroyed the model. This rather surprising result can be understood by employing a very simple analytic model, namely the linear shallow-water equations for a barotropic ocean over flat bottom on an infinite f -plane. Assuming perfect measurements of sea surface height one can obtain an equation of the form (4.5) for the state deviation from truth. The system matrix is diagonalizable with one zero eigenvalue representing the stationary geostrophic mode and a complex conjugate pair referring to inertia-gravity waves. By actually evaluating the right-hand side of equation (4.6) one can indeed prove general divergence of the nudging approach (4.7-4.8) for the simple linear system. Even for the initial state already being in geostrophic balance, due to the effect of the nudging term distorting the original eigenmodes, energy seeps into the altered inertia-gravity modes where it can accumulate and finally destroy the model.

A more encouraging result can be obtained from a second approach to construct the nudging terms. Alternatively to (4.7-4.8) one can nudge the total velocity (and not only the geostrophic part) towards the “observed” geostrophic velocity:

$$\frac{\partial u}{\partial t} = \text{“physics”} + \mu \left(-\frac{g}{f} \frac{\partial \eta^{obs}}{\partial y} - u \right) \quad (4.9)$$

$$\frac{\partial v}{\partial t} = \text{“physics”} + \mu \left(\frac{g}{f} \frac{\partial \eta^{obs}}{\partial x} - v \right) . \quad (4.10)$$

In this case, test calculations with the GFDL model did not indicate any divergence problems. Again it is straightforward to analyze the linear barotropic shallow-water model and to show that the above choice of a diagonal matrix \mathbf{M} indeed guarantees convergence of the assimilation procedure when perfect data are assumed. This is in sharp contrast to the convergence behaviour of the first approach (4.7-4.8) and clearly underlines the importance of a proper structure of the nudging matrix \mathbf{M} .

Unfortunately, for complex non-linear systems like high-resolution ocean circulation models it is not yet possible to give a definite rule for the structure of the nudging matrix \mathbf{M} . However, it is fair to say that, to be cautious, one should try using a positive diagonal matrix whenever possible. Throughout this work I shall therefore use the reduced form of the extended nudging approach,

$$\frac{\partial \mathbf{x}}{\partial t} = \text{"physics"} + \mathbf{M} (\mathbf{H} \mathbf{y}^{obs} - \mathbf{x}) , \quad (4.11)$$

with \mathbf{M} being a positive diagonal matrix and $\mathbf{H} = \mathbf{M}^{-1}\mathbf{N}$. By writing $\mathbf{M} = \text{diag}(\mu_1, \dots, \mu_N)$, it becomes evident that each component of the state vector obeys a differential equation essentially of the form (4.1). This assures that $\mathbf{M} (\mathbf{H} \mathbf{y}^{obs} - \mathbf{x})$ always acts as a damping term on the model-data misfit and does not generate any spurious wave modes by introducing some non-physical coupling between different components of the state vector.

It is important to note that the extended nudging approach (4.11) requires the observations to be projected onto the model state space, what in general will be an underdetermined inverse problem. For the latter reason, more refined assimilation schemes like the Kalman filter always use the well-defined forward projection of the model state onto the observations. However, as shown above, it is the continuous action of the nudging term disturbing the physics of the model, that makes the special form (4.11) necessary.

To proceed, some guiding principles shall be presented for introducing time dependency of the diagonal nudging matrix \mathbf{M} , which is important if the observations are intermittent rather than continuous in time. Subsequently, I will investigate possible choices of the matrix \mathbf{H} , needed to project the vector of observations, \mathbf{y}^{obs} , onto the state space of the model.

4.3 Intermittent Nudging

For eddy-resolving primitive-equation models of the ocean, like the one to be used in this study, the length of the individual model timesteps is of the order of one hour, which for most oceanographic data sets including mapped altimeter data is some orders of magnitude smaller than typical intervals between successive observations. In this case the basic idea of the nudging approach is that around each observation time a number of timesteps are used to nudge the model state towards the observed variables, while the dynamical and thermodynamical processes of the model can spread out the information introduced by the observations, so that a balanced state, consistent with the model dynamics can be obtained (HOKE and ANTHES, 1976). This dynamic initialization is of particular importance for primitive-equation models, and in practise the balanced state is often assumed to be reached, when the fast inertia-gravity waves, excited for example by observations not in geostrophic balance, have decayed to a level acceptable for a reasonable model evolution.

By acting continuously over a number of timesteps, the nudging term for a single observation \mathbf{x}^{obs} behaves like a damping term for all those oscillatory modes that, during this time span, run through more than half a cycle. While being desirable for the fast inertia-gravity waves excited by unbalanced imperfect data, one usually would not want to dampen, for example, quasi-geostrophic motions. For this reason it is common to introduce some time-dependence of the nudging parameter $\mu = \mu(t)$, to essentially switch off the action of the nudging term after some time span longer than typical periods of inertia-gravity waves, but short when compared with quasi-geostrophic timescales. If the interval between successive observations is large enough, one thereby obtains a well defined time window during which the model state is nudged towards a single observation vector \mathbf{y}^{obs} .

Using a Gaussian time-dependence of the nudging parameter, HOLLAND and MALANOTTE-RIZZOLI (1989) compared this local (in time) data insertion with continuous assimilation of the same observations interpolated linearly in time. Their results suggest that an assimilation procedure local in time is likely to be more effective than the simple linear interpolation approach, the former in addition having the conceptual advantage of allowing the model to evolve freely except for the time when the nudging is turned on. Taking the above considerations into account we shall express the time-dependence of the nudging parameter to be used throughout this work in form of a

Gaussian centered at the observation time t_0 ,

$$\mu(t) = \mu_0 \frac{1}{\sqrt{\pi\tau}} \exp \left[- \left(\frac{t - t_0}{\tau} \right)^2 \right] , \quad (4.12)$$

for $t_0 - \Delta T/2 \leq t < t_0 + \Delta T/2$, where ΔT is the period between successive observations. In this form the nudging constant μ_0 represents the total weight of the nudging term, which, following the smooth shape of the Gaussian, is “spread out” over a number of model timesteps (Fig. 4.1). The width of the Gaussian is controlled by the parameter τ

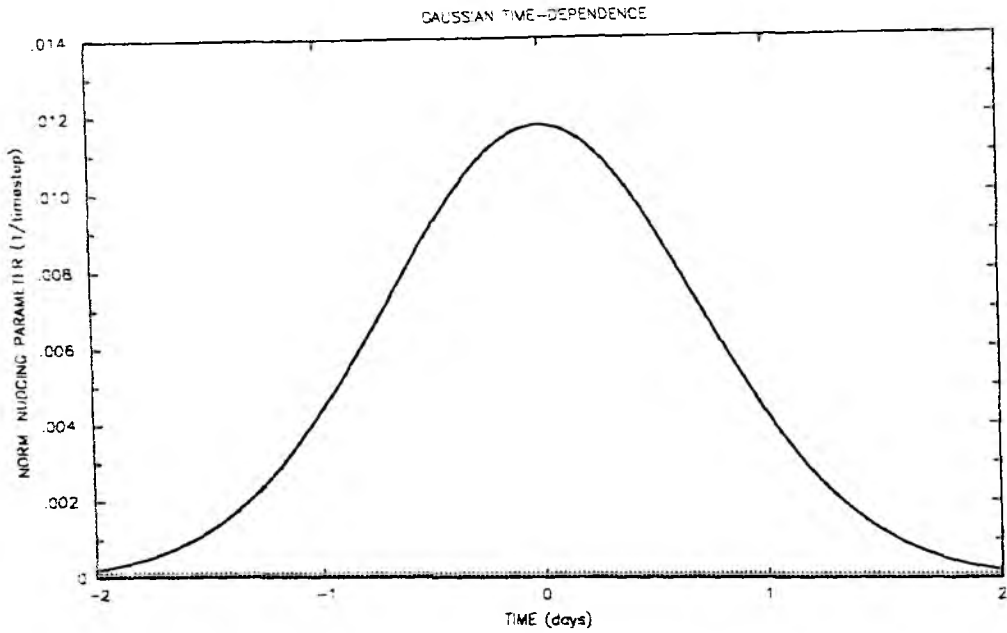


Figure 4.1: The Gaussian time-dependence of the normalized nudging parameter $\mu(t)/\mu_0$ for a typical length of model timesteps of 0.5 h and $\tau = 48$ timesteps ≈ 1 day. Individual model time steps are indicated by the small tickmarks on the time axis.

and, as argued above, should be short when compared with timescales of motions one does not want to dampen. For eddy-resolving ocean models, we therefore suggest the upper limit for τ be of the order of one day. If, on the other hand, a primitive-equation model is used and the observations are not properly balanced, τ has to be large enough to avoid too strong radiation of inertia-gravity waves emerging from the adjustment process by which the model dynamically reacts to the data. Since the timescale of this adjustment is given by the inverse of the Coriolis parameter f (e.g., GILL, 1982), we

conclude that for our objective $\tau = 1$ day is a reasonable value as long as one cannot guarantee that the projected data $\mathbf{x}^{obs} = \mathbf{H}\mathbf{y}^{obs}$ are geostrophically balanced.

For completeness we remark that for quasi-geostrophic models, with inertia-gravity waves filtered out, it was actually shown by DOMBROWSKY and DEMEY (1992) that the nudging window can be reduced to one timestep (i.e., re-initialization) if a rather complete “pseudo” observation \mathbf{x}^{obs} of the full state vector is assimilated. However, if the observations comprise only very few components of the state vector, what for example is the case for along-track assimilation of altimeter data as performed by VERRON *et al.* (1992), some finite width τ of the nudging window can be appropriate even when using a quasi-geostrophic model. Because of the sparseness of the data, a simple re-initialization of the observed state variables could quickly be swamped by the “inertia” of the large portion of the model, that is not directly affected by the observations. Allowing the information given by the observations to enter the model over a number of timesteps, the model dynamics can contribute to establish a complete state that is consistent with the sparse data. However, it will turn out in the following sections that this dynamical interpolation is rather inefficient for primitive-equation models, both in the vertical direction (section 4.4) as well as in the horizontal (section 6.2). For the assimilation of altimeter data into such a model it will therefore be fundamental to develop some appropriate operator \mathbf{H} that provides a rather complete “pseudo” observation \mathbf{x}^{obs} prior to the actual dynamical assimilation procedure. Hence, the only purpose of using a finite width of the nudging window in this study is the damping of inertia-gravity waves. Actually, the re-initialization approach to be developed in chapter 5 is just based on the construction of a properly balanced “pseudo” observation \mathbf{x}^{obs} , and consequently allows a reduction of the nudging window to a single model timestep. This is in close correspondence to the findings of MALANOTTE-RIZZOLI *et al.*, (1989), that for a primitive-equation model with a rigid lid a geostrophically balanced initialization is sufficient to ensure a smooth evolution of the model fields.

The remaining parameter to be estimated is the constant μ_0 , that essentially represents the integral weight of the nudging term corresponding to a single observation vector \mathbf{x}^{obs} . Before turning towards this problem in section 4.5, we shall first investigate the question of an appropriate choice of the projection operator \mathbf{H} for the objective of our study.

4.4 The Projection Operator \mathbf{H}

Any assimilation scheme based on the original nudging approach (4.1) first of all implies that all the observed variables are prognostic variables of the model. As was shown in section 4.2, it is nevertheless possible to extend the original scheme to nudging of non-prognostic variables, if only some relationship to prognostic model variables exists and a careful choice of the nudging matrix \mathbf{M} is made. Cautiously restricting ourselves to a positive diagonal nudging matrix, the formulation

$$\frac{\partial \mathbf{x}}{\partial t} = \text{"physics"} + \mathbf{M} (\mathbf{H} \mathbf{y}^{obs} - \mathbf{x}) \quad (4.13)$$

of the extended nudging scheme was recommended. In general, the choice of the projection operator \mathbf{H} is far from being unambiguous. In the case of satellite altimetry it is not obvious how to relate, for example, observed sea surface height to the salinity field of a numerical model. Strictly speaking, it simply is not possible to statically invert the surface observation for a unique complete three-dimensional state of the ocean. Usually a large nullspace remains, and in fact it is one of the driving motives of data assimilation to impose dynamical constraints in order to reduce this nullspace. (See also the Appendix on observability.)

A series of so-called identical twin experiments was carried out in order to test possible choices of the projection operator \mathbf{H} for their suitability for assimilating altimeter data into an eddy-resolving primitive-equation model. For computational economy simplified box configurations, described in sections 3.2 and 3.3, rather than the expensive CME model of the North Atlantic Ocean were used for these experiments. First, a control run – serving as substitute for the real world – was performed and the time evolution of its state vector and the corresponding surface pressure field stored. These surface pressure maps were subsequently assimilated into the same numerical model, but started from completely different initial conditions. An objective criterion, quantifying the ability of the assimilation scheme to let the assimilation run converge towards “reality”, is given by the difference between the state vectors of assimilation run and control run, respectively.

For the moment we shall concentrate on the problem of vertical transfer of surface information and defer the discussion of the pay-off between horizontal and temporal resolution implied by the satellite’s orbit schedule to chapter 6. Here, complete and perfect maps of absolute sea surface height are assimilated every five days, allowing the model to evolve freely between the well-defined Gaussian nudging windows of width $\tau =$

1 day. As argued in the previous section, this is a physically reasonable choice, also confirmed by a number of test calculations using different values for τ . So far we do not know how best to choose the integral intensity of the nudging represented by the nudging coefficient μ_0 . However, since the sea surface height observations as provided by the control run are free of errors, we are tempted to proceed from a rather large value for μ_0 . In most of the experiments to be described below, the standard value of this coefficient is therefore set to $\mu_0 = 1$, i.e., just enough to completely replace the model state \mathbf{x} by the one inferred from the observations, $\mathbf{x}^{obs} = \mathbf{H}\mathbf{y}^{obs}$. This corresponds to an amplitude of the nudging parameter of $\mu(t = t_0) = \mu_0/(\sqrt{\pi}\tau) \simeq (2 \text{ days})^{-1}$. While this value is adequate for perfect observations, it will later be reduced to take into account both errors in the surface data and also inaccuracies of the projection scheme \mathbf{H} . Generally, the model integration was stopped after 16 assimilation cycles at day 82, when most of the experiments seemed to have approximately reached their asymptotic error level. A small number of integrations performed for another 80 days confirmed this choice.

Probably the most natural approach to constructing \mathbf{H} is to ignore the nullspace and restrict the projection $\mathbf{H}\mathbf{y}^{obs}$ to those components of the state vector that can be reliably deduced from the data. Considering altimetry, it is thus possible, with some assumption of geostrophic balance, to directly infer geostrophic surface current velocities from observations of sea surface height. In this respect it seems appropriate to nudge towards the “observed” geostrophic surface velocity in the prognostic equations for the surface velocity components only, while leaving the prognostic equations of velocity components at deeper levels as well as of temperature and salinity unchanged. This approach of nudging the surface currents only and leaving the vertical transfer of information to the internal model dynamics, has been applied quite successfully in the context of quasi-geostrophic models (e.g., HOLLAND and MALANOTTE-RIZZOLI, 1989; VERRON, 1992), and also for an adiabatic layered primitive-equation model (HAINES *et al.*, 1993). We shall therefore start by applying the very same method to the GFDL model.

In a first series of experiments, hereafter named UV1L (nudging of u and v at the 1st level only), the operator \mathbf{H} projects the complete and perfect maps of sea surface height onto the velocity components u and v at the uppermost level of the model by employing the geostrophic assumption. (Observations south of 5°N are neglected.) Before being assimilated into the model, this “observed” geostrophic velocity is corrected for the Ekman velocity, which, assuming the Ekman transport to be confined to the surface layer, is easily computed for the simplified box model configurations being

UV1L Experiments					
		4 levels		21 levels	
		$\mu_0 = 1$	$\mu_0 = 10$	$\mu_0 = 1$	$\mu_0 = 10$
velocity	(surface)	31	75	14	38
	(middle)	30	69	12	38
	(deep)	12	34	6	15
temperature	(basin)	21	42	6	22

Table 4.1: Total percentage decrease of rms error after 16 assimilation cycles (82 days). “Middle” refers to the second level (225 m) in the four-level model and to the fifth level (174 m) in the 21-level model, “deep” to the fourth (2350 m) and 15th level (1800 m), respectively.

forced with steady winds. There are no “pseudo” observations for any other components of the state vector, i.e., for temperature, salinity and deep velocities. Accordingly, we set $\mu_0 \neq 0$ in the momentum equations for the first level only.

For the four-level model described in chapter 3.3, the results are rather disappointing when the standard nudging coefficient $\mu_0 = 1$ is used. The convergence of the assimilation run towards the control run is considerably worse than one would expect from the above mentioned previous studies using quasi-geostrophic models. The final reduction of the rms error at day 82 relative to the respective values when no data are assimilated, is presented in the first column of Table 4.1. However, more encouraging results are obtained with very hard nudging ($\mu_0 = 10$, corresponding to $\mu(t_0) \simeq (0.2 \text{ days})^{-1}$). As shown in Figure 4.3, at day 90 the error levels of the velocity field at the first and second model level have virtually reached their asymptotic limit, while the errors of the bottom currents and also that of temperature still decrease. The reduction of the rms error after 16 assimilation cycles is now 75% (69%) for the velocity at the first (second) level (Table 4.1). This is close to the error reduction of about 80% obtained by HOLLAND and MALANOTTE-RIZZOLI (1989) for the streamfunctions of the upper and middle layers of their three-layer quasi-geostrophic model. This result, which is also consistent with the findings of HAINES *et al.* (1993), suggests that – at least under certain circumstances – it is indeed appropriate to assimilate altimeter data into primitive-equation models by simply nudging the surface velocity.

To proceed, essentially the same UV1L experiments were repeated with the 21-level model described in section 3.2. Again standard ($\mu_0 = 1$) and hard ($\mu_0 = 10$) nudging was applied, however, with significantly less success as in the four-level case! The evolution of the rms error is shown in Figure 4.4, with the final reduction in the rms error given in the third and fourth column of Table 4.1. Note particularly that

even with very hard nudging the model is not able to appropriately recover the surface velocity field. Apparently, nudging just the surface current field becomes less effective as the vertical resolution of the numerical model is increased. This phenomenon is believed to be due to three reasons:

- First, in a primitive equation model surface information propagates into the ocean interior via the action of internal gravity waves. As noted earlier, the timescale of geostrophic Rossby adjustment processes is f^{-1} on a spatial scale of the Rossby radius of the corresponding vertical modes. For models with a rather coarse vertical resolution, this apparently is still rapid enough to obtain results comparable to that of quasi-geostrophic models that mediate surface information by the instantaneous process of vortex stretching (realized at each timestep through the inversion of the potential vorticity for the streamfunction). However, increasing the vertical resolution adversely affects this rapid adjustment, as more and more vertical modes come into play, with ever decreasing speed of propagation and increasing distortions due to the numerical discretization imposed by the finite horizontal grid spacing.
- Second, the input of momentum via the nudging terms is proportional to the thickness of the surface layer. Hence, the observations have a stronger and more persistent impact when a model with a low vertical resolution is used. This is most important if the geostrophic adjustment can not be completed within the nudging time window (e.g., because of a too large number of vertical modes). As can be deduced from Figure 4.4 (curve *C*), the rms error of a thin (35 m) surface layer quickly increases after each assimilation cycle. This behaviour is due to the little inertia relative to the remainder of the water column with a total depth of 3500 m. In contrast, the surface-layer thickness of the four-level model is 100 m to be compared with an overall depth of 3600 m.
- Third, geostrophic adjustment to lowest order conserves potential vorticity on isopycnals. In fact, HAINES *et al.* (1993) used this property to explain the success of nudging the surface currents in both quasi-geostrophic and also primitive-equation layer models. Their argument crucially depends on the implicit assumption that the variability in the potential vorticity field be contained within the uppermost layer of the model. As discussed by HAINES (1991), this is indeed a rather good approximation for layer models that do not allow for isopycnal surfacing (and thereby actually impose limitations on the vertical resolution). However, for the primitive-equation 21-level model used in this study, the basin averaged modulus of vertical correlations between potential vorticity of the first and fifth model level (174 m) amounts to 0.56, what is significantly different from zero! Vertically correlated anomalies in the potential

vorticity field will in particular originate from episodic water mass formation in regions where the depth of the mixed layer exceeds that of the surface layer of the model. An example of this phenomenon was given by COX (1987), who used the potential vorticity field to track the movement of intermittently formed mode waters.

The above results clearly indicate that for our purposes it is advisable to look for a more effective projection operator H . Given the high vertical coherence of mesoscale oceanic variability (e.g., DEMEY and ROBINSON, 1987) it is generally believed that the information about anomalies of geostrophic surface current velocities, as contained in altimeter observations, reflects more than just the surface conditions. To investigate this matter in greater detail, first the vertical correlations between the modelled horizontal velocity components at the surface level and those at depth $z(k)$, $\text{Cor}(u_k, u_{surf})$, $\text{Cor}(v_k, v_{surf})$, and also the cross-correlations $\text{Cor}(u_k, v_{surf})$ and $\text{Cor}(v_k, u_{surf})$, were calculated for each column of model grid points from a climatology run extending over three model years. For the simplified models used in the identical twin experiments the windstress is constant in time as is the corresponding Ekman transport. However, the CME model uses monthly winds, and one has to make sure that the unsteady ageostrophic Ekman velocity, which is not observable by altimetry, does not enter the correlation coefficients. Since the depth δ_E of the Ekman layer, given by $\delta_E = (2K_M/f)^{1/2}$, for the vertical viscosity coefficient of $K_M = 10 \text{ cm}^2 \text{ s}^{-1}$ is less than 10 m, it is assumed, that the Ekman transport is contained completely in the first layer of the model with a thickness of 35 m. For the CME model therefore not the first level, but the second level velocity components are referred to as surface velocities in the computation of the velocity correlations.

It has been shown by CHENEY *et al.* (1984) that in the northwest Atlantic Ocean the depth of the 15°C isotherm is significantly correlated with sea surface height. Here we have to deal with a numerical model formulated on fixed depth levels. It was therefore decided to further compute correlations between sea surface height and temperature at any depth level k , $\text{Cor}(T_k, \eta)$, as well as between sea surface height and salinity, $\text{Cor}(S_k, \eta)$, in addition to the velocity correlation coefficients introduced above. In general, components of horizontal velocity were found to be strongly correlated quasi-barotropically (with correlation coefficients exceeding 0.6) to the respective surface velocity components down to a depth of 500 m (i.e., $k = 9$), decreasing with depth, while the absolute values of the cross-correlation terms on average did not exceed 0.2 and hence were neglected in the following. A typical example for vertical profiles of correlation coefficients $\text{Cor}(u_k, u_{surf})$ is shown for a section along 30°N in Figure 4.2 *a*. In contrast, profiles of $\text{Cor}(T_k, \eta)$ exhibit a somewhat different vertical structure. They

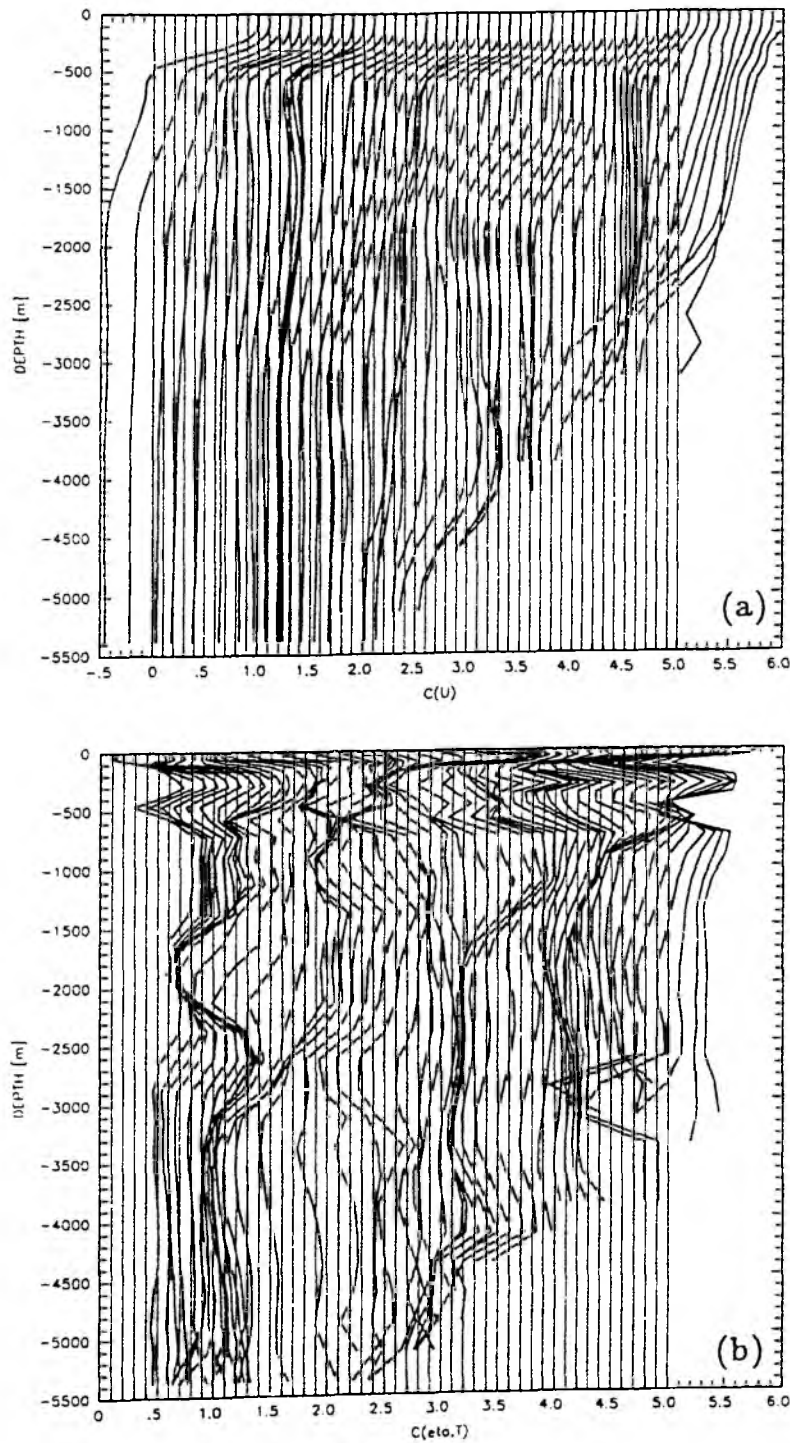


Figure 4.2: Profiles of vertical correlation coefficients between (a) zonal component of surface velocity and zonal component of deep velocities, and (b) surface elevation and temperature. Profiles for all model gridpoints along 30°N from 60°W to 40°W are drawn, each with an offset of 0.1. (Profiles from the North Atlantic Ocean CME model.)

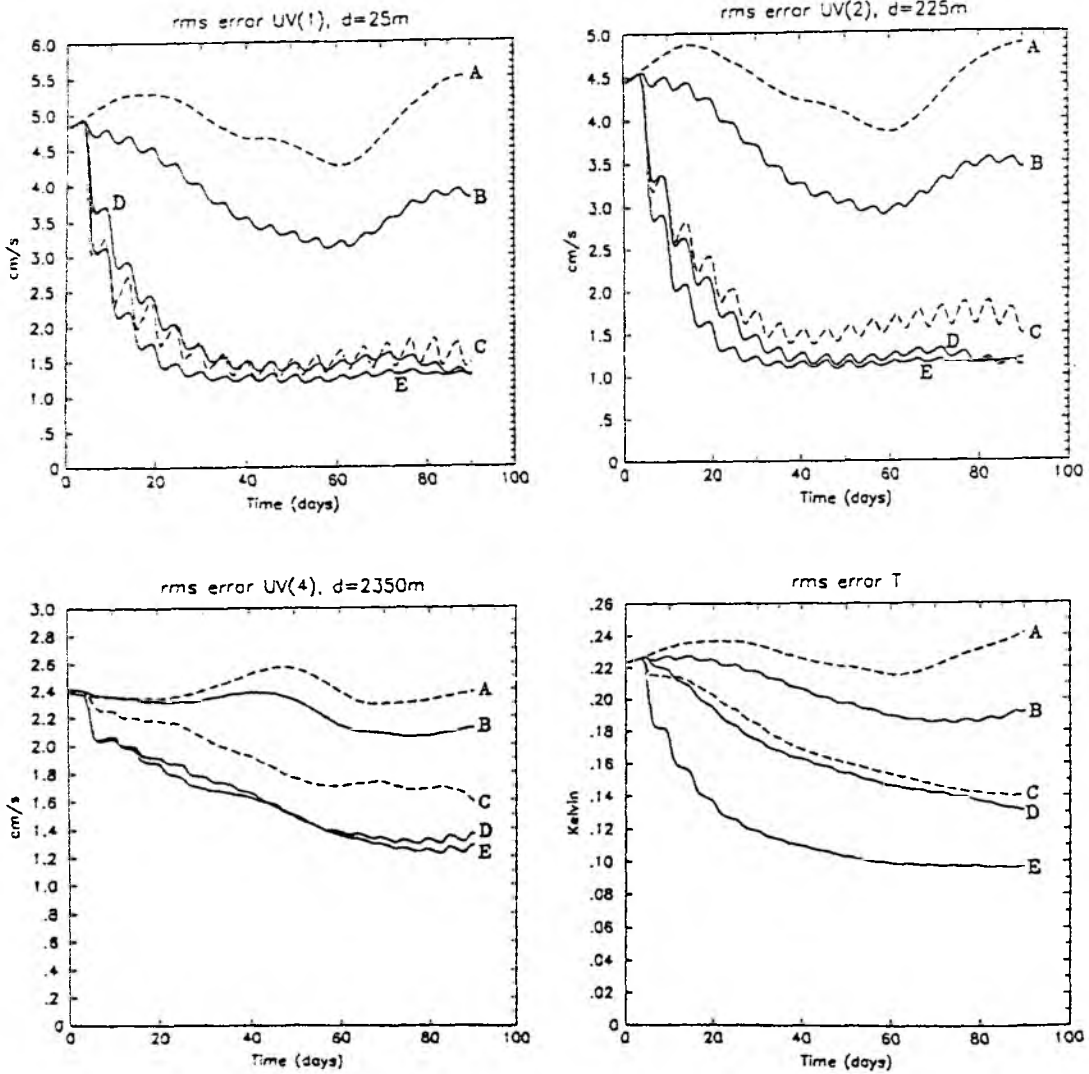


Figure 4.3: Time evolution of the rms differences between control run and assimilation run for the 4-level model. Values for velocities are shown for selected depth levels (25 m, 225 m, 2350 m), while the rms difference for temperature is computed over the entire basin. Curves (A) denote the reference run when no data are assimilated, (B) and (C) refer to nudging of the first-layer velocity only (experiment UV1L) with nudging timescales $(2 \text{ days})^{-1}$ and $(0.2 \text{ days})^{-1}$, respectively. Curves (D) refer to experiment LRUV, using linear regression coefficients for u and v , and (E) to experiment LRUVTS, using linear regression coefficients for all state variables, with maximum nudging intensities of $(2 \text{ days})^{-1}$.

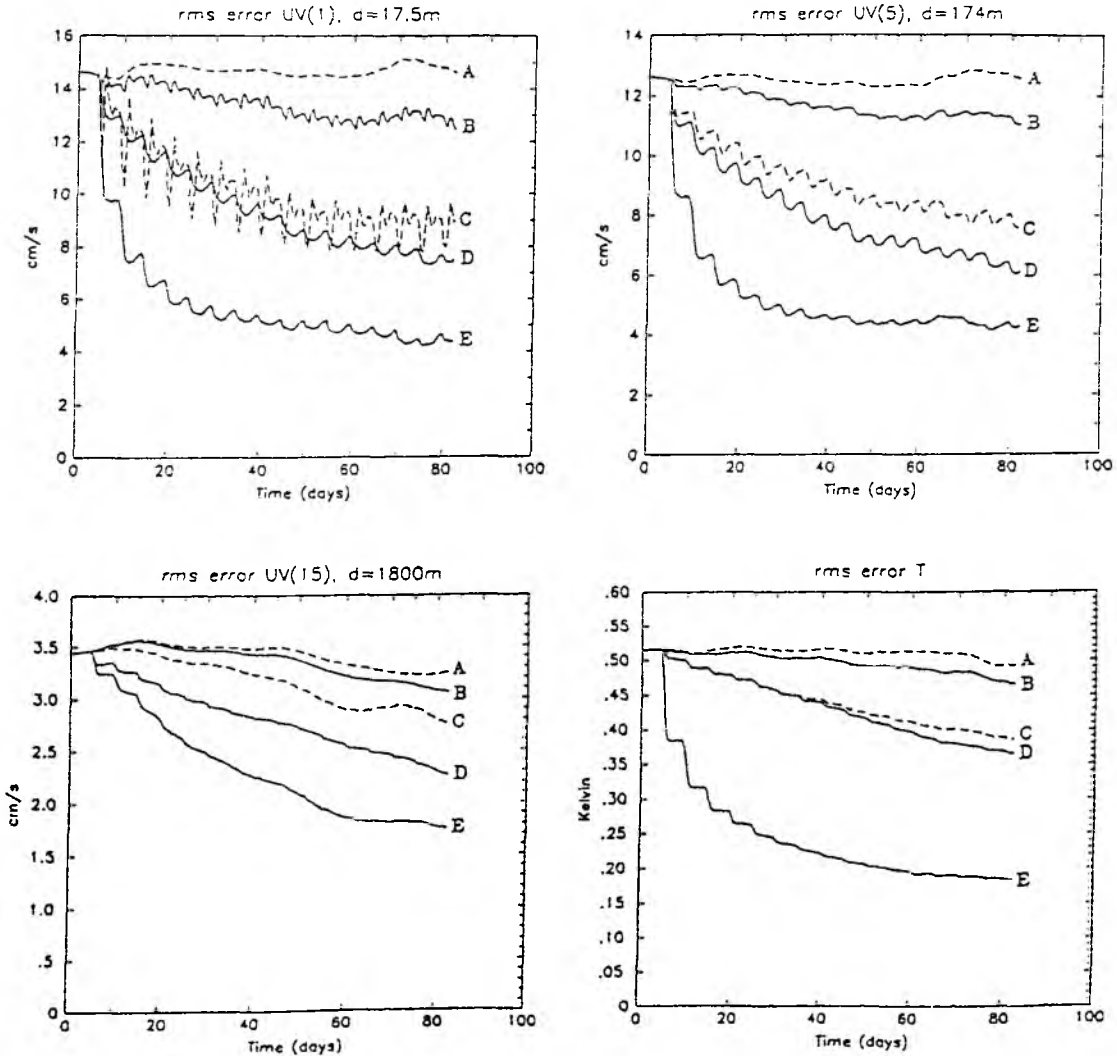


Figure 4.4: Time evolution of the rms differences between control run and assimilation run for the 21-level model. Values for velocities are shown for selected depth levels (17.5 m, 174 m, 1800 m), while the rms difference for temperature is computed over the entire basin. Curves (A) denote the reference run when no data are assimilated, (B) and (C) refer to nudging of the first-layer velocity only (experiment UV1L) with nudging timescales $(2 \text{ days})^{-1}$ and $(0.2 \text{ days})^{-1}$, respectively. Curves (D) refer to experiment LRUV, using linear regression coefficients for u and v , and (E) to experiment LRUVTS, using linear regression coefficients for all state variables, with maximum nudging intensities of $(2 \text{ days})^{-1}$.

often indicate a maximum at the depth of the thermocline, but generally their structure is much less regular than that of the velocity correlation profiles, with strong variations from grid point to grid point (see Figure 4.2*b*). We finally note that when calculating the correlation coefficients, no attempt has been made to separate meso-scale fluctuations from seasonal or even interannual variability, because we expect the satellite altimeter to detect oceanic variability on all these time scales.

The correlation coefficients thus computed can be further used to relate any state variable of the model to either sea surface height or surface velocity, both – assuming geostrophy – being directly observable by altimetry. As noted previously, the dynamics of the ocean are too complex to allow for an exact static inversion of the surface pressure field for the full three-dimensional state of the model. For this reason, we shall resort to a statistical concept that ideally should approximate the dominant mode of the unknown exact inversion, treating the remaining degrees of freedom as noise. We start from the hypothesis that, in the mean, some linear relationships between observed surface variables and the remaining components of the state vector exist. For example, the zonal velocity component u and the potential temperature T at depth level k may be written as

$$u_k = \langle u_k \rangle + R(u_k) \delta u_{surf} + \varepsilon(u_k) \quad (4.14)$$

$$T_k = \langle T_k \rangle + R(T_k) \delta \eta + \varepsilon(T_k) \quad , \quad (4.15)$$

where $\langle \dots \rangle$ denotes the climatological mean, $\delta u_{surf} = (u_{surf} - \langle u_{surf} \rangle)$, and the quantity ε is the actual deviation from the postulated linear relationship. One would like to determine the regression coefficients R in such a way that on average the deviations ε are minimized. It is well known that if the distribution of the ε were Gaussian with zero mean, then the maximum likelihood estimate for the coefficients R would be given by the least squares fit,

$$R(u_k) = \frac{\langle \delta u_k \delta u_{surf} \rangle}{\langle (\delta u_{surf})^2 \rangle} = \text{Cor}(u_k, u_{surf}) \left(\frac{\text{Var}(u_k)}{\text{Var}(u_{surf})} \right)^{1/2} \quad (4.16)$$

$$R(T_k) = \frac{\langle \delta T_k \delta \eta \rangle}{\langle (\delta \eta)^2 \rangle} = \text{Cor}(T_k, \eta) \left(\frac{\text{Var}(T_k)}{\text{Var}(\eta)} \right)^{1/2} \quad , \quad (4.17)$$

minimizing the variance of the deviations, $\langle \varepsilon^2 \rangle$, from the estimated values

$$u_k^{est} = \langle u_k \rangle + R(u_k) \delta u_{surf} \quad (4.18)$$

$$T_k^{est} = \langle T_k \rangle + R(T_k) \delta \eta \quad . \quad (4.19)$$

Similar equations can be set up for v_k^{est} and S_k^{est} . This linear regression scheme is a rather simple method for projecting the surface observations vertically into the deep

ocean, with the regression coefficients representing the mean vertical profile of deviations from the climatological mean state. Because no horizontal correlations are used, this projection scheme is local in observation space: The estimate of any state variable is based on a single surface data point. On the one hand this may neglect some information present in the data, on the other hand it limits the storage requirements for the projection operator H to just twice the dimension of the model's state vector. (One vector is required for storing the climatological mean values, $\langle \dots \rangle$, a second one for the corresponding regression coefficients, R .) A sketch of this projection scheme is displayed in Figure 4.5.

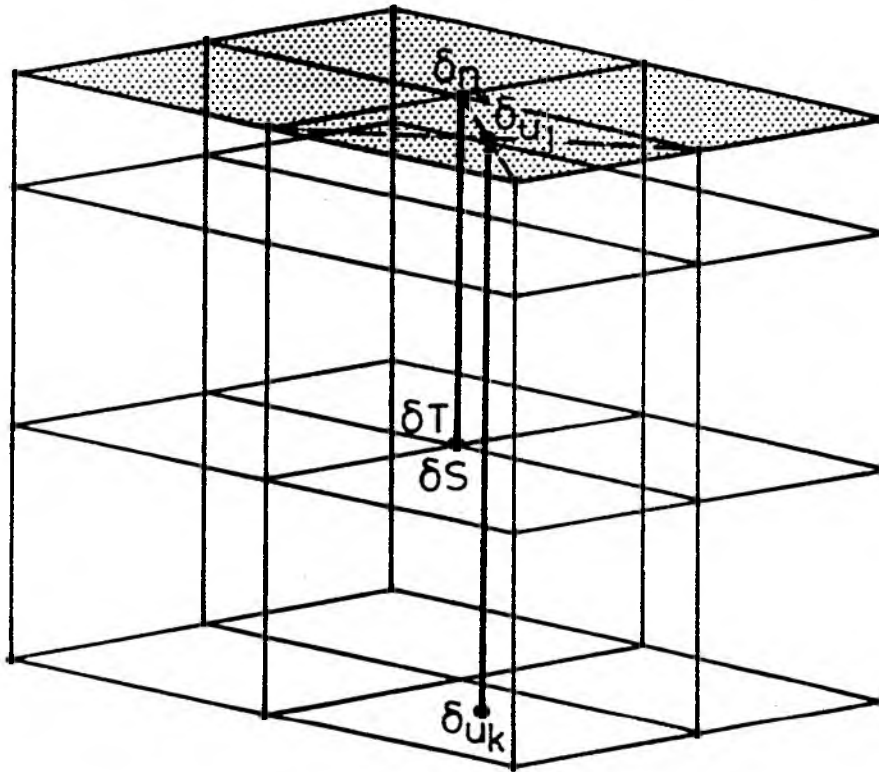


Figure 4.5: Symbolical sketch of the vertical projection scheme that linearly relates temperature T and salinity S to anomalies $\delta\eta$ of sea surface height, and velocity components u and v to anomalies of the surface values u_{surf} and v_{surf} .

While the above introduced linear regression method straightforwardly provides some estimate \mathbf{x}^{est} of the complete state vector based solely on surface observations, it is still appropriate to check the relevance of this estimate. First, the underlying statistical presumptions have to be examined, namely the proposed linear relationship (4.14-4.15) with normally distributed deviations ε around the straight line. Actually, the data used for computing correlation and regression parameters are taken from the model climatology (42 states sampled at intervals of 26 days over a period of three years). Consequently the deviations are not random numbers, but emerge from the deterministic dynamics of the ocean model. After all, scatter diagrams based on this climatological data set generally indicate that the individual points may indeed be considered as being independent from each other. Typical scatter plots are shown for the position 30°N, 30°W in Figure 4.6. However, at some positions the interannual variability – or even a systematic model trend – clearly dominates mesoscale fluctuations. In such cases values obtained from a sampling interval of 26 days can not be taken as independent. A rather special example is depicted in Figure 4.7. At this station the easily recoverable sequence of points indicate a warming of the CME model of about 2°C at 900 m over three model years, that is associated with a local sea level rise of roughly 10 cm! Nevertheless, we shall for the moment assume that all of the $N = 42$ deviations $\varepsilon(T_k)$ are independent and, moreover, come from the same (unknown) normal distribution. It can then be shown (e.g., DRAPER and SMITH, 1981) that the quantity

$$t = \sqrt{\frac{N-2}{1 - \text{Cor}^2(T_k, \eta)}} \text{Cor}(T_k, \eta) \quad (4.20)$$

is a random number of a t -distribution with $N - 2$ degrees of freedom. This can be further exploited to test the null hypothesis that no simple linear relationship of the form (4.14-4.15) exists, i.e., $R(u_k) = 0$ and $R(T_k) = 0$. The horizontally basin averaged significance level for rejecting this null hypothesis is plotted for the velocity components u_k and v_k as well as for the state variables T_k and S_k as a function of depth $z(k)$ in Figure 4.8. Generally, the vertical projection scheme based on the statistical linear regression approach is most reliable for the velocity components above the main thermocline, where we also expect most of the variability to occur. This may give some confidence for using the linear regression approach (4.18) to directly infer estimates about the deep velocity field from the observed surface currents. On the other hand, the significance of the regression coefficients for temperature and salinity turns out to be rather low over the whole depth of the water column, and it may be questionable whether they shall be used at all to infer subsurface information from surface data.

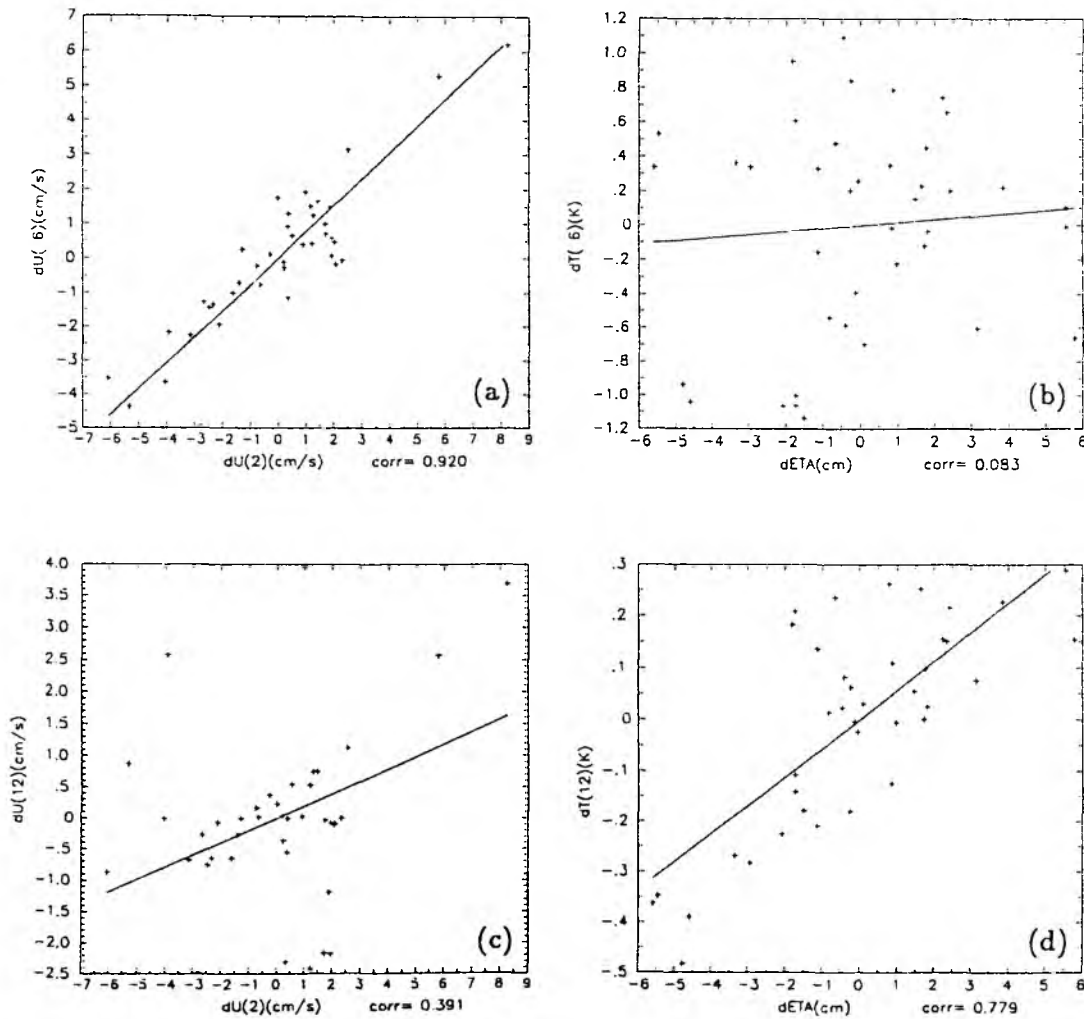


Figure 4.6: Scatter diagram for a station at 30°N , 30°W of (a) zonal velocity at 230 m against the zonal velocity at the surface, (b) temperature at 230 m against surface elevation, (c) zonal velocity at 900 m against the zonal velocity at the surface, and (d) temperature at 900 m against surface elevation. The least-square regression line is also shown, with the corresponding correlation value printed below each frame. All data are taken at 26 day intervals from the North Atlantic Ocean CME model.

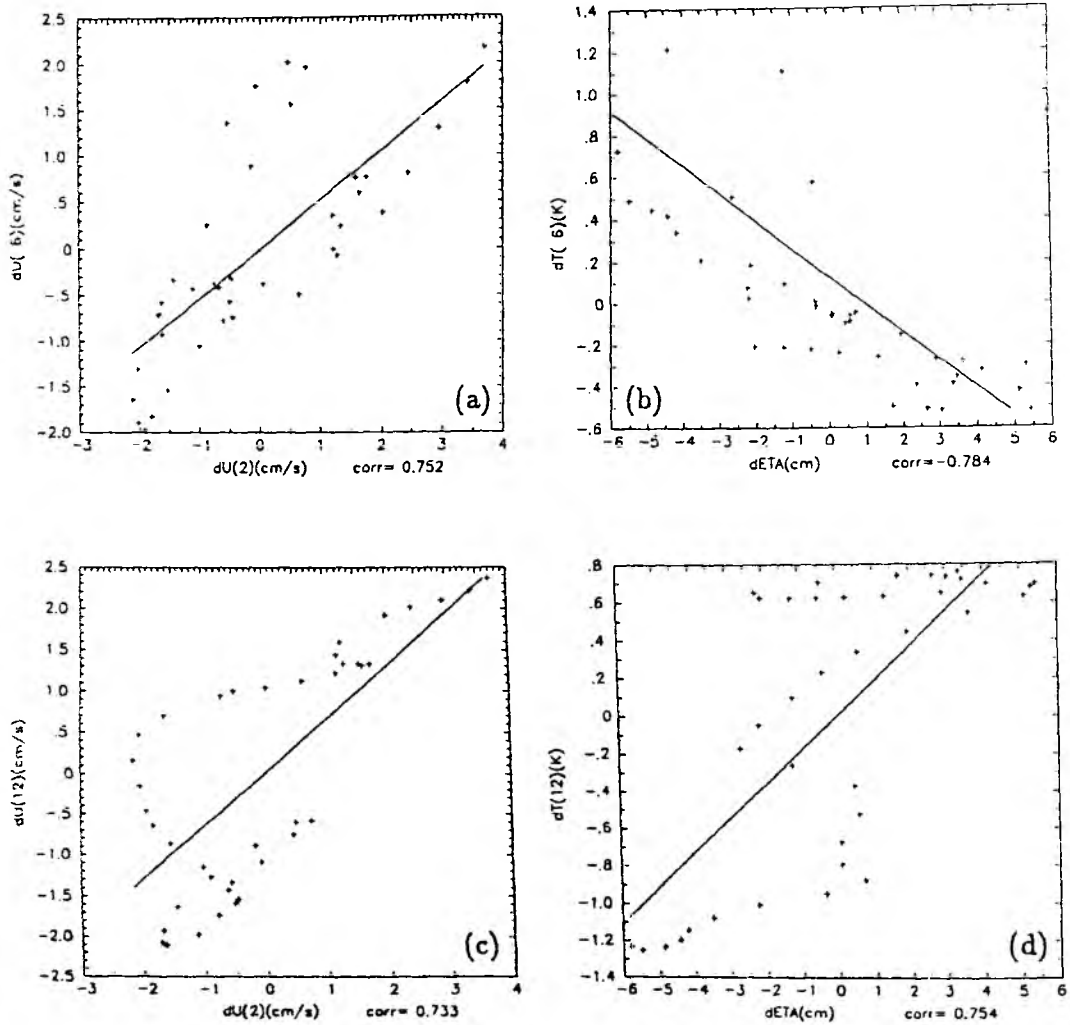


Figure 4.7: Scatter diagram for a station at 40°N , 30°W of (a) zonal velocity at 230 m against the zonal velocity at the surface, (b) temperature at 230 m against surface elevation, (c) zonal velocity at 900 m against the zonal velocity at the surface, and (d) temperature at 900 m against surface elevation. The least-square regression line is also shown, with the corresponding correlation value printed below each frame. All data are taken at 26 day intervals from the North Atlantic Ocean CME model.

To further investigate this matter, an additional series of identical twin experiments was performed. In the first experiment, named LRUV (Linear Regression for u and v), only the regression coefficients relating subsurface velocity components to the observed surface velocity are used to provide pseudo observations $(u, v)_k^{obs} \equiv (u, v)_k^{est}$, whereas the second experiment, LRUVTS (Linear Regression for u , v , T , and S), also exploits the generally weaker vertical correlations between temperature/salinity and sea surface height to obtain a complete state vector of pseudo observations, $\mathbf{x}^{obs} \equiv \mathbf{x}^{est}$.

Taking into account that the simple linear regression method explains only the portion Cor^2 of the actual variance, e.g., $\text{Var}(T_k^{est}) = \text{Cor}^2(T_k, \eta) \text{Var}(T_k)$, the elements of the diagonal nudging matrix \mathbf{M} are downweighted by the square of the corresponding correlation coefficient, and we set $\mu_0 = \text{Cor}^2$. Like in the previous experiments the width of the nudging window is chosen as $\tau = 1$ day, hence $\mu_0 = \text{Cor}^2$ corresponds to $\mu(t_0) \simeq \text{Cor}^2 \cdot (2 \text{ days})^{-1}$. As being discussed in greater detail in section 4.6, the choice $\mu_0 = \text{Cor}^2$ ensures that the model state is explicitly nudged only towards those pseudo observations that can be reliably estimated by the simple vertical projection scheme.

The evolution of the rms errors for the two twin experiments LRUV and LRUVTS is shown by curves (D) and (E) in figures 4.3 and 4.4. Even though the use of the linear regression method for the velocity field already leads to a much more rapid and effective convergence than observed for the UV1L experiments with nudging of the surface layer velocity only, there is still some improvement gained when the correlations between temperature/salinity and surface elevations are also exploited in the vertical projection scheme. Not surprisingly, this is most pronounced for the temperature field, that directly absorbs the additional information provided by the temperature correlations. Nevertheless there is also considerable difference in the convergence of the velocity field, that in the LRUV experiment tends to geostrophically re-adjust to the unchanged density field. Consequently, some part of the velocity information may eventually be rejected. This emphasizes the importance of directly projecting the surface observations onto the density structure, as for example is implemented in experiment LRUVTS.

To conclude, we note that as far as identical twin experiments are concerned, combining statistical vertical projection schemes with a simple dynamical assimilation method can quite effectively nudge the model state close to the (simulated) reality. With regard to limited storage space, any projection method considered in this work can only exploit vertical correlations between the prognostic model fields and directly observable surface variables. This approach may be viewed as a statistically optimal “feature model” (in analogy to the empirical feature models suggested by ROBINSON *et al.*,

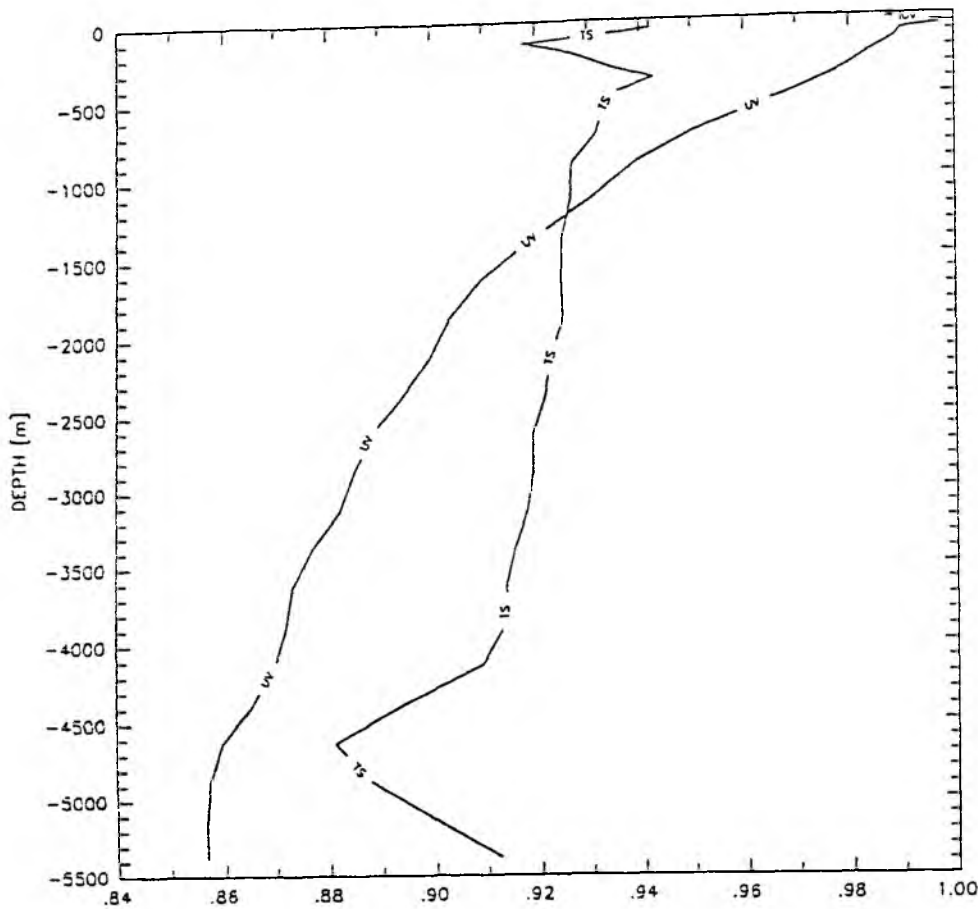


Figure 4.8: Significance level for rejecting the null hypothesis $R = 0$, horizontally averaged over the entire model domain, as a function of depth $z(k)$. The curve labeled (UV) refers to the regression coefficients for the velocity components, $R(u_k)$ and $R(v_k)$, the one labeled (TS) refers to $R(T_k)$ and $R(S_k)$. (Curves shown are from the North Atlantic Ocean CME model.)

1988), with the strength of the nudging depending on the reliability of the estimated features. Best results are obtained when all variables of the state vector are directly related to the surface observations, i.e., when all variables are included in the feature model. Clearly, applying the nudging term to the surface velocities only is not very efficient if only the vertical resolution of the model is fine enough. In order to reach reasonable convergence times one should exploit any information available to directly project the surface data into the ocean interior. This conclusion is expected to hold for a free-surface model as well, as simply nudging the prognostic surface elevation corresponds to nudging the surface velocities in a model with a rigid lid.

4.5 Nudging versus Optimal Interpolation

After having gained some information about the appropriate structure of the nudging matrix \mathbf{M} as well as of its time-dependence in sections 4.2 and 4.3, we still have to estimate optimal values for the absolute size of its matrix elements, namely the nudging coefficients μ_0 , referring to the integral weight of the Gaussian (4.12). First of all it is evident that the size of μ_0 should depend on the accuracy of the (pseudo) observations as compared with the confidence we have in the state provided by the model forecast. For very accurate observations one would like large values of μ_0 to force the model state towards these observations. When, on the other hand, one trusts the model forecast more than the data, one would prefer a rather small weight of the nudging term to avoid severe distortions of the model physics.

Besides this rather qualitative consensus, presently no established concept exists, how to best determine the nudging coefficients. For some very simple linear systems it is possible to determine an optimal μ_0 by requiring the nudging procedure to converge towards the result of the statistically optimal adjoint assimilation method. A different approach uses the adjoint technique for the model with the nudging terms included to numerically compute the optimal nudging coefficients μ_0 by taking the μ_0 themselves as control parameters (ZOU *et al.*, 1992). However, because this estimate of the optimal μ_0 is a numerical one only, it is rather difficult to generalize the results obtained for a specific example. The practical use of this method has still to be proven, especially with regard to such systems which, for computational reasons, do not allow an implementation of the adjoint technique.

Here, we will develop an alternative approach to estimate the parameters μ_0 in the context of the extended nudging approach by stressing its analogy to sequential optimal interpolation. Concentrating on a proper choice of the elements of the nudging matrix \mathbf{M} , we shall for the moment neglect any effects due to a particular projection matrix \mathbf{H} and simply refer to $\mathbf{x}^{obs}(t_0)$ as (pseudo) observations provided at time t_0 . The time-discretized form of the nudging equations then reduces to

$$\mathbf{x}(t) = \text{“physical evolution”} + \Delta t \mathbf{M} \left(\mathbf{x}^{obs}(t_0) - \mathbf{x}(t - \Delta t) \right) , \quad (4.21)$$

where, with regard to ease implementation into existing codes and also for numerical stability reasons (VERRON, 1992), a simple Euler forward discretization scheme has been applied.

While the width τ of the nudging window can be chosen on dynamical considerations (see section 4.3), it is only the parameter μ_0 that remains to be estimated,

essentially representing the integral weight of the nudging term corresponding to a single observation vector $\mathbf{x}^{obs}(t_0)$. Since the nudging term contains the state vector which may change during the Gaussian time window, partly as an effect of the nudging itself, the meaning of the integral weight μ_0 is not immediately clear. For the moment we shall therefore consider a slightly modified version of the nudging equation (4.21): Let the model run in the pure forecast mode, with no nudging applied, to compute a model forecast $\mathbf{x}^f(t_0)$ at the observation time t_0 . Then use the same initial conditions at $t = t_0 - \Delta T/2$, where again ΔT is the period between observations, to start the assimilation run with the nudging term now containing the stored model forecast, $\mathbf{x}^f(t_0)$,

$$\mathbf{x}(t) = \text{"physical evolution"} + \Delta t \mathbf{M} (\mathbf{x}^{obs}(t_0) - \mathbf{x}^f(t_0)) \quad . \quad (4.22)$$

The state emerging from this assimilation cycle can then in turn be taken as initial condition for a subsequent model forecast run. The use of the fixed state $\mathbf{x}^f(t_0)$ over the entire time window permits direct comparison of this nudging scheme with statistical optimal interpolation. If the (pseudo) observational state vector \mathbf{x}^{obs} is not strongly unbalanced, nudging can be regarded as statistical interpolation "spread out in time". At least for linear systems this suggests the optimal choice

$$\mu_0 = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_{obs}^2} \quad , \quad (4.23)$$

where for simplicity the errors of the observation and those of the model forecast are assumed to be completely uncorrelated, with known covariances σ_{obs}^2 , and σ_f^2 , respectively. While the above relation may serve as a sound statistical basis for the modified approach, we still have to investigate its relevance for the usual nudging procedure (4.21). If, for example, the model state is very different from the observations, we expect the impact of the original procedure to be weaker than that of the modified one. This is due to the ability of the state vector $\mathbf{x}(t)$ to adjust to the observations $\mathbf{x}^{obs}(t_0)$ already during the actual nudging process, thus decreasing the integral effect of the nudging term in equation (4.21).

This expectation is indeed confirmed by the results of identical twin experiments, performed with the North Atlantic box model configuration described in section 3.2. For both nudging schemes, (4.21) and (4.22), complete maps of sea surface height were assimilated every five days, using the same vertical projection scheme and nudging coefficients as for the LRUVTS experiment. The evolution of the rms error is displayed in Figure 4.9. Clearly, the faster initial convergence of the modified approach is evident during the first few assimilation cycles. However, it is interesting to see that both

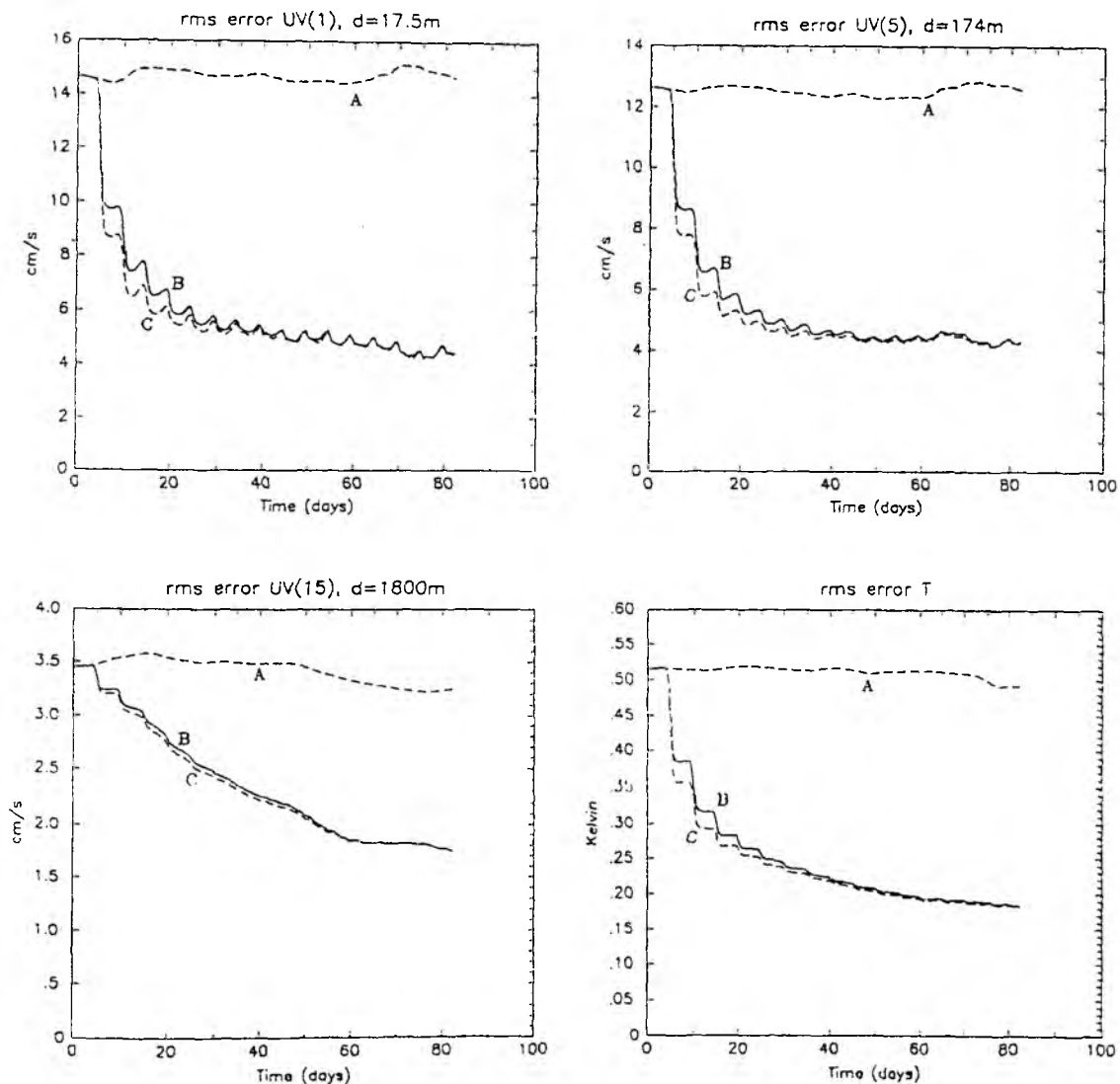


Figure 4.9: Time evolution of the rms differences between control run and assimilation run for the 21-level model. Values for velocities are shown for selected depth levels (17.5 m, 174 m, 1800 m), while rms differences for temperature are computed over the entire basin. Curves (A) denote the reference run when no data are assimilated, (B) refers to the original nudging approach (4.21), and (C) is the result of the modified nudging method (4.22).



methods finally lead to the same asymptotic error level. Moreover, the actual model fields at the end of the assimilation experiments are virtually identical. This result confirms that it can indeed be reasonable to consider nudging as optimal interpolation “spread out in time”.

4.6 Choice of the Nudging Matrix M

In this section, a sensible choice of the nudging matrix M , only based on statistical principles, will be recommended. As argued in section 4.2, a positive diagonal matrix is cautiously used throughout this study. The Gaussian time dependence of its diagonal elements, the nudging parameters $\mu(t)$, was introduced in section 4.3, with the width of the time window generally set to $\tau = 1$ day. As seen in the previous section it seems appropriate to consider nudging of (pseudo) observations of the complete state vector simply as statistical interpolation spread over a number of timesteps, thereby ensuring a smooth evolution of the model state and avoiding the generation of large amounts of gravity waves by unbalanced re-initialization shocks. The analogy between nudging and optimal interpolation provides some indication for a proper choice of the integral intensity of the nudging, represented by the parameter μ_0 for the integral weight of the nudging term (see Eq. 4.12).

To simplify the analysis it will be assumed that observational and model forecast errors are uncorrelated. Furthermore we neglect any correlations between errors of different variables, which reduces the error covariance matrices to their diagonal elements. Each nudging parameter μ_0 can then in principle be computed from the simple scalar relation

$$\mu_0 = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_{obs}^2} , \quad (4.24)$$

where σ_f^2 denotes the error variance of the model forecast, and σ_{obs}^2 is the error variance of the (pseudo) observation. For example, let us consider the nudging parameter for the state variable T_k : Assuming the correlation coefficient used in the vertical projection scheme to be correct, the error variance of the estimated pseudo observation T_k^{est} is given by

$$\sigma^2(T_k^{est}) = R^2(T_k)\sigma^2(\eta^{obs}) + [1 - \text{Cor}^2(T_k, \eta)] \text{Var}(T_k) . \quad (4.25)$$

The first term on the right is the downward extension of the actual observational error and the second term is the error due to the imperfect vertical correlation ($\text{Cor}^2 \neq 1$) between temperature and surface elevation. For the normalized error variance this

yields

$$\frac{\sigma^2(T_k^{est})}{\text{Var}(T_k)} = 1 - \text{Cor}^2(T_k, \eta) \left[1 - \frac{\sigma^2(\eta^{obs})}{\text{Var}(\eta)} \right] . \quad (4.26)$$

Expression (4.24) can now be rewritten as

$$\mu_0(T_k) = \frac{C_f^2(T_k)}{C_f^2(T_k) + 1 - \text{Cor}^2(T_k, \eta) [1 - \sigma^2(\eta^{obs})/\text{Var}(\eta)]} , \quad (4.27)$$

where the $C_f^2(T_k)$ is the unknown error variance of the model forecast normalized by the variance of the temperature field. While estimates of the normalized error variance of the surface height data can be drawn from the satellite measurements and from the mapping routine used to interpolate the observations onto the model grid (chapter 6), the sparsity of oceanographic data presently does not really allow for an accurate determination of the model forecast error.

Although it is well known that the CME model has systematic rather than just statistical deficiencies, we shall continue by assuming its forecast error to be of statistical nature. However, a rather conservative choice is made by setting $C_f^2 = 1$. To ensure that this conservative estimate of the forecast error does not allow extremely inaccurate pseudo observations to be weighted too much, only the first order term of the series expansion of expression (4.27) about $\text{Cor}^2(T_k, \eta) [1 - \sigma^2(\eta^{obs})/\text{Var}(\eta)] = 1$ (i.e., perfect observations) is retained. The chosen nudging coefficient $\mu_0(T_k)$ then becomes

$$\mu_0(T_k) = \text{Cor}^2(T_k, \eta) [1 - \sigma^2(\eta^{obs})/\text{Var}(\eta)] , \quad (4.28)$$

and accordingly

$$\mu_0(u_k) = \text{Cor}^2(u_k, u_{surf}) [1 - \sigma^2(u_{surf}^{obs})/\text{Var}(u_{surf})] . \quad (4.29)$$

In the identical twin experiments considered so far, the observations are free of errors. Hence $\sigma^2(\eta^{obs}) = \sigma^2(u_{surf}^{obs}) = 0$ and $\mu_0 = \text{Cor}^2$ is used in the experiments LRUV and LRUVTS described above. On the other hand, the observational error variances have to be taken into account, for both the maps-versus-tracks studies of chapter 6 and, of course, for the assimilation experiments using GEOSAT data. Note that $\text{Var}(\eta)$ and $\text{Var}(u_{surf})$ are the variances of the model fields. As was shown by STAMMER and BÖNING (1992), the CME model greatly underestimates the rms sea surface variability (see also chapter 7 of this work). To avoid this inaccuracy of the model statistics leading to negative nudging coefficients, it has been decided to replace $\text{Var}(\eta)$ by $\text{Var}(\eta^{obs})$. An estimate of $r = \sigma^2(\eta^{obs})/\text{Var}(\eta^{obs})$ is provided then by the mapping routine, which is based on an algorithm developed by DEMEY and MÉNARD (1989), used to interpolate

the tracked altimeter data onto the model grid (chapter 6). With some knowledge of $\sigma^2(\eta^{obs})$ and of the correlation radius used in the mapping routine, one could in principle also estimate the error variance of the “observed” geostrophic surface velocities, that are computed from the mapped altimeter measurements using finite differences. However, in order to ensure that always $\left[1 - \sigma^2(u_{surf}^{obs})/\text{Var}(u_{surf})\right] \geq 0$, and hence $\mu_0(u) \geq 0$, it was somewhat pragmatically decided to replace this term by $[1 - r]^2$, with $r = \sigma^2(\eta^{obs})/\text{Var}(\eta^{obs})$. The final choice of the nudging parameters then reads

$$\mu_0(T_k) = \text{Cor}^2(T_k, \eta) [1 - r] \quad , \quad (4.30)$$

and

$$\mu_0(u_k) = \text{Cor}^2(u_k, u_{surf}) [1 - r]^2 \quad , \quad (4.31)$$

with equivalent formulations for $\mu_0(S_k)$ and $\mu_0(v_k)$.

The above expressions fix the integral weights of the nudging coefficients with Gaussian time dependence as described by equation (4.12). By taking account of the accuracy of the individual pseudo observations, the elements of the nudging matrix are interrelated with the statistical vertical projection scheme developed in section 4.4. Consistently with previous considerations, the model state is not altered ($\mu_0 = 0$) if either there is no surface information ($r = 1$) or the vertical correlation is zero. In the opposite case of perfect surface observations ($r = 0$) combined with perfect vertical correlations, the model variables are essentially replaced by the observations ($\mu_0 = 1$). Although still being affected by the particular formulation of forecast and observational error estimates, the above introduced model for the nudging parameters finally achieves the purpose of setting up a nudging matrix \mathbf{M} without any free tuning parameters.

4.7 Summary of the Extended Nudging Approach

In the final section of this chapter, the actual implementation of the extended nudging approach shall be demonstrated by explicitly writing down the modifications of the model equations, introduced by the nudging terms. For brevity the original terms due to the model physics are just included symbolically. While in the CME model there are prognostic equations for horizontal velocity components u and v , temperature T and salinity S at each model level k (chapter 3), only the equations for u_k and T_k are displayed exemplarily:

$$\frac{\partial u_k}{\partial t} = \text{"physics"} + \mu_{u_k}(t) [\langle u_k \rangle + R(u_k) \delta u_{surf}^{obs} - u_k] \quad (4.32)$$

$$\frac{\partial T_k}{\partial t} = \text{"physics"} + \mu_{T_k}(t) [\langle T_k \rangle + R(T_k) \delta \eta^{obs} - T_k] \quad (4.33)$$

The regression parameters $R(u_k)$ and $R(T_k)$ are given by equations (4.16) and (4.17), $\langle \dots \rangle$ is the climatological mean, and the nudging coefficients are chosen as

$$\mu_{u_k}(t) = \text{Cor}^2(u_k, u_{surf}) [1 - r]^2 \frac{1}{\sqrt{\pi\tau}} \exp \left[- \left(\frac{t - t_0}{\tau} \right)^2 \right] , \quad (4.34)$$

and

$$\mu_{T_k}(t) = \text{Cor}^2(T_k, \eta) [1 - r] \frac{1}{\sqrt{\pi\tau}} \exp \left[- \left(\frac{t - t_0}{\tau} \right)^2 \right] , \quad (4.35)$$

where $r = \sigma^2(\eta^{obs})/\text{Var}(\eta^{obs})$ and $t_0 - \Delta T/2 \leq t < t_0 + \Delta T/2$, with ΔT being the period between successive observations. As argued in section 4.3 the above formulations shall be generally used with $\tau = 1$ day. Test calculations using different values for the width of the nudging window gave very similar results for the range $0.5 \text{ days} < \tau < 2 \text{ days}$. Larger values for τ led to a considerable decrease of the kinetic energy level due to damping of high-frequency motions. When, on the other hand, τ was reduced to a single timestep, large amounts of small-scale internal gravity noise was generated by the not properly balanced disturbance that efficiently opposed a smooth evolution of the model.

Although the explicit form of the extended nudging approach retains only the *anomalies* δu_{surf}^{obs} and $\delta \eta^{obs}$, this does not imply that knowledge of the *mean* sea surface height (which due to the large geoid error can not satisfactorily be measured by altimetry) is not required by this assimilation method. Actually, the statistical linear regression method relies on the first and second moments (i.e., mean and covariance) of the model fields being consistent with the assimilated GEOSAT data. All these

statistical moments, including the mean sea surface height, are taken from the model and are assumed to accurately reflect real oceanic conditions.

As a matter of routine, the modulus of the nudging term has been compared with the modulus of the total physical forcing. Thereby it was found that in the prognostic momentum equations the forcing due to the nudging term in general did not exceed 10% of the forcing due to the model physics, while it could reach 50% of the physical forcing in the tracer equations. Hence, at each single timestep it is still the physics that dominates the evolution of the state vector. However, because the nudging term acts more or less in the same direction over a number of timesteps, it can have a substantial integral effect, with the small individual weights of the nudging terms ensuring that the state vector always remains practically consistent with the model physics.

Twin experiment results clearly indicated that some vertical projection of the surface restricted altimeter data onto the complete set of model state variables was appropriate. In the present chapter a statistical regression method was applied to extrapolate the actual observations of sea surface height anomalies into the ocean interior. It is obvious that this method relies heavily on accurate statistics. The regression scheme used further implies that the assimilation of SSH anomalies will essentially conserve the climatology of the model. While being quite reasonable for identical twin experiments, this property is certainly not always desirable when real data are assimilated into a numerical model with possibly systematic deficiencies. An alternative assimilation method that depends much less on statistical relationships shall be developed in the following chapter. Eventually, the performance of both assimilation schemes will be compared, first within the identical twin approach, and finally under more realistic conditions by assimilating GEOSAT data.

Chapter 5

Re-Initialization

5.1 Beyond the Extended Nudging Approach

Despite its high degree of realism, unprecedented by any other basin-scale eddy-resolving ocean circulation model, the CME configuration of the North Atlantic Ocean still exhibits considerable systematic deficiencies. All the simulations performed so far have shown a rather unrealistic flow pattern in the Gulf Stream separation region, and the representation of mean frontal structures and eddy variability as associated with the North Atlantic Current (NAC) and the Azores Current (AC) is not really satisfactory (BECKMANN *et al.*, 1994*a,b*). Quite generally, the intensity of mesoscale variability is found to be considerably underestimated (STAMMER and BÖNING, 1992), a phenomenon that is most pronounced in the eastern basin.

Actually, all these model deficiencies provide some stimulus to the effort of assimilating observations in order to improve the model's performance. A first simple assimilation method adapted for the assimilation of satellite altimeter data into this model is the extended nudging approach developed in the previous chapter. From the results of identical twin experiments, and also for dynamical considerations, it was concluded that for the assimilation of such observations into a high vertical resolution model to be efficient, some direct vertical extrapolation of the surface data into the ocean interior was required. A computationally feasible projection method was constructed on the basis of local vertical correlations between directly observable surface properties and the actual state variables of the model. However, because of the lack of oceanographic data, the required correlation coefficients could not be taken from reality, but instead were computed from a climatology run of the numerical model

itself. Applying these statistics to actual GEOSAT data essentially implies that the model statistics closely resemble that of the real ocean. Having in mind the significant underestimation of mesoscale variability by the model, this is reason for some concern.

First, it is likely that the contribution of the seasonal fluctuations, which the model simulates quite realistically (BÖNING *et al.*, 1991; BÖNING and HERRMANN, 1994), to the correlation coefficients is overestimated. In particular, this will apply to the temperature field, which for most regions of the upper ocean is expected to show the largest seasonal signal compared to mesoscale variability. While this relative overestimation of seasonal effects could in principle be eliminated by using different regression coefficients for different frequency bands, there remains a second concern that raises more serious problems: There is absolutely no guarantee that an extrapolation of the simple linear regression scheme (4.18-4.19) to such altimeter measurements that considerably exceed the range of sea level fluctuations simulated by the model will give physically reasonable results. Indeed, this will turn out to be a serious problem especially for estimates of temperature and salinity. Because of the physical constraints on these variables to closely match the regional distribution of water masses, any linear relationship between water mass properties and sea surface height will be rather local in phase space and one should be extremely reluctant to further extrapolate it. In this respect, the regression coefficients of the horizontal velocity components, which may be viewed as representing the average superposition of different vertical modes, are believed to depend much less on the typical amplitudes of the surface anomalies. Furthermore, it has already been shown that even when the statistics of the CME model are assumed to be perfect, a linear relationship between surface currents and velocity components down to the main thermocline is more likely to hold than such a relation between surface elevation and temperature or salinity (Figure 4.8).

On account of the above considerations we shall try to provide an alternative vertical projection scheme that relies less on the rather uncertain model statistics, but instead has a sounder physical basis and hence should be more robust to systematic errors of the model. Priority is given to replacing the statistical linear regression method for temperature and salinity by some physically more appealing dynamical method, that geostrophically couples changes in the density field to changes in the velocity field, the latter still being estimated statistically. As will turn out later – thereby retrospectively justifying the particular implementation of this alternative approach –, even the approximate obedience of this balance condition will allow substitution of the smooth and rather continuous nudging of the prognostic variables by an instantaneous re-initialization of the state vector. At observation time the model is stopped, the

data are used to update the model forecast state, whereupon the model is restarted from this newly computed, so-called analyzed state vector. This will have important consequences as the constraint on the weighting matrix M being diagonal, essentially due to the permanent action of the non-physical nudging term (section 4.2), can now be dropped.

Note that the re-initialization method to be developed in the following section is in some respect the very opposite of the update assimilation technique suggested by MELLOR and EZER (1991), who used a vertical linear regression scheme (with regression coefficients provided by the model statistics) to estimate temperature and salinity of the model from sea surface height measurements, whereas changes in the velocity field were computed subsequently by requiring geostrophic balance with the statistically inferred updates of the model density field. Although the differences between both methods will become most apparent, when real data are combined with model statistics, it will be shown in section 5.4 that even for an identical twin experiment, i.e., the perfect statistics case, the assimilation scheme of MELLOR and EZER is less effective than the re-initialization approach proposed in this study.

5.2 The Method

Let η^f be the surface height field of the model forecast at observation time t_0 at which measurements of anomalies of the surface elevation, $\delta\eta^{obs}$, are available. Using the statistical regression coefficients, $R(u)$, for the velocity field (Eq. 4.16), one can project the model-data misfit in the surface height gradients (i.e., the geostrophic surface velocities) downwards and obtain the estimated velocity errors, $R(u_k) \left[(-g/f) \partial_y \delta\eta^{obs} - (-g/f) \partial_y \delta\eta^f \right]$, of the model forecast velocity u_k^f at each depth level k . As previously, $\delta\eta^f = \eta^f - \langle \eta \rangle$, where $\langle \eta \rangle$ is the climatological mean surface height taken from the model. In this respect, the model-data misfit is assumed to have the same vertical structure as a typical anomaly of the velocity field. Its projection can then be directly used to update the velocity field of the model forecast, yielding the so-called analyzed estimate

$$u_k^a = u_k^f + \mu_0(u_k) R(u_k) \left[-\frac{g}{f} \frac{\partial \delta\eta^{obs}}{\partial y} + \frac{g}{f} \frac{\partial \delta\eta^f}{\partial y} \right] \quad (5.1)$$

with some weighting coefficient $\mu_0(u_k)$. Note that exactly the same method for estimating the model-data misfit led to divergence of the extended nudging scheme! As discussed in detail in section 4.2, this behaviour was due to the permanent action of the non-diagonal nudging term distorting the model physics, and in fact was a main reason for choosing a positive diagonal nudging matrix thereafter. Consequently, the extended nudging scheme was essentially based on pseudo observations of the complete state vector itself that could be compared with the actual model state. In contrast, the instantaneous re-initialization requires estimated (“pseudo”) model-data differences only. This implies that, if $\nabla \delta\eta^f = \nabla \delta\eta^{obs}$, the analyzed deep velocity components u_k^a would simply remain the u_k^f of the model forecast. Actually, this quite reasonable feature is not ensured by the extended nudging scheme (4.32)!

The second conceptual advantage of the updating method (5.1) over the extended nudging scheme is that there is no implicit constraint of leaving the climatological mean of any model variable unchanged under the assimilation of sea surface height anomalies. After all, for computational reasons both methods approximate optimal interpolation between model forecast and observations by neglecting any non-local covariances that, for example, would show up in the equivalent equation of the statistically optimal Kalman filter. Here, the optimal weighting matrices shall be approximated by scalar weighting parameters μ_0 that take into account the accuracy of the surface data and also of the vertical extrapolation scheme. Along the same arguments that in section 4.6 led to the choice of the integral weights used for the extended nudging scheme, we again set $\mu_0(u_k) = \text{Cor}^2(u_k, u_{surf})[1 - r]^2$, where $r = \sigma^2(\eta^{obs})/\text{Var}(\eta^{obs})$ is the estimated relative error variance of the altimeter data as provided by the mapping routine (see chapter 6).

Equation (5.1) uses the extrapolated model-data misfit to update the model forecast. Together with the corresponding relation for the meridional velocity components this provides an analyzed estimate of the complete velocity field. However, in order to avoid vigorous excitation of internal/inertia gravity waves emerging from the Rossby adjustment process that would try to re-adjust the new velocity field back to the density field of the model forecast and possibly lead to rejection of the data, some concurrent update of the temperature and salinity variables is also required. As an alternative to using statistical correlations between sea surface height and water mass properties we shall proceed by exploiting dynamical principles to compute a density anomaly that is consistent with the estimated velocity misfit \hat{u}_k and \hat{v}_k , defined by

$$\hat{u}_k = R(u_k) \left[-\frac{g}{f} \frac{\partial \delta\eta^{obs}}{\partial y} + \frac{g}{f} \frac{\partial \delta\eta^f}{\partial y} \right] , \quad (5.2)$$

and

$$\hat{v}_k = R(v_k) \left[\frac{g}{f} \frac{\partial \delta \eta^{obs}}{\partial x} - \frac{g}{f} \frac{\partial \delta \eta^f}{\partial x} \right] \quad (5.3)$$

For the time and space scales that can be detected by satellite altimetry, geostrophy is an appropriate balance criterion that interrelates velocity changes with changes in the density field. Since the statistically estimated velocity deviations in general will not be non-divergent, any attempt to geostrophically balance these fields will to some extent have to rely on simplifying assumptions.

By combining the hydrostatic and geostrophic assumptions, the thermal wind equation relates vertical shear in the horizontal velocity field to horizontal gradients of the *in situ* density field. This relation is linear and can be applied to the estimated errors \hat{u} , \hat{v} , and $\hat{\rho}$ of the model forecast directly, yielding

$$\frac{\partial \hat{u}}{\partial z} \equiv -\frac{g}{f} \frac{\partial R(u)}{\partial z} \frac{\partial}{\partial y} [\delta \eta^{obs} - \delta \eta^f] = \frac{g}{f \rho_0} \frac{\partial \hat{\rho}}{\partial y} \quad (5.4)$$

$$\frac{\partial \hat{v}}{\partial z} \equiv \frac{g}{f} \frac{\partial R(v)}{\partial z} \frac{\partial}{\partial x} [\delta \eta^{obs} - \delta \eta^f] = -\frac{g}{f \rho_0} \frac{\partial \hat{\rho}}{\partial x}, \quad (5.5)$$

where $\rho_0 = \rho_0(z)$ refers to the background density profile. As a first approximation the vertical regression coefficients $R(u)$ shall be assumed to be horizontally homogeneous. This is in fact supported by the corresponding vertical correlations exemplarily shown in Figure 4.2. With the spatial derivatives ∂_x and ∂_y acting on the differences between two much stronger undulating instantaneous fields, it is therefore reasonable to set

$$\int \frac{\partial R(u)}{\partial z} \frac{\partial}{\partial y} [\delta \eta^{obs} - \delta \eta^f] dy \simeq \frac{\partial R(u)}{\partial z} [\delta \eta^{obs} - \delta \eta^f] + \text{const}(x, z). \quad (5.6)$$

This can be related to the integral of the right hand side of equation (5.4), resulting in an expression for the density increment required to balance the deviations in the zonal velocity geostrophically:

$$\hat{\rho} \simeq -\rho_0 \frac{\partial R(u)}{\partial z} [\delta \eta^{obs} - \delta \eta^f] + \text{const}(x, z). \quad (5.7)$$

An alternative expression for the same quantity can be obtained from the analogous balance condition for the meridional velocities:

$$\hat{\rho} \simeq -\rho_0 \frac{\partial R(v)}{\partial z} [\delta \eta^{obs} - \delta \eta^f] + \text{const}(y, z). \quad (5.8)$$

If the estimated velocity errors were non-divergent, one could in principle uniquely determine the integration constants except for an unknown function of the vertical coordinate z . However, the statistical regression coefficients do not automatically satisfy this criterion. For this reason there is no self-consistent set of integration constants, and consequently a unique density increment $\hat{\rho}$ does not exist.

To eliminate this ambiguity, a second approximation is introduced: The anomaly patterns are assumed to be horizontally isotropic with their vertical structure given by the local average of zonal and meridional regression profiles. This readily implies $\text{const}(x, z) = \text{const}(y, z) \equiv \text{const}(z)$. Proposing further that for a perfect forecast of the sea surface elevation, i.e., $\delta\eta^f = \delta\eta^{obs}$, the density field should not be changed, finally leads to the choice $\text{const}(z) = 0$. The resulting estimate for the density increment which approximately balances the estimated velocity changes is

$$\hat{\rho} = -\rho_0 \frac{\partial}{\partial z} \left(\frac{R(u) + R(v)}{2} \right) \cdot [\delta\eta^{obs} - \delta\eta^f] . \quad (5.9)$$

It is interesting to note that this relationship between density changes and surface height anomalies is again a local one that can be evaluated for each water column individually! Of course, such a local relationship could only be obtained by employing two simplifying assumptions, namely the vertical structure of deviations from the model forecast being horizontally homogeneous and isotropic.

To gain some information about the actual degree of geostrophic balance achieved by this procedure, the output density increments $\hat{\rho}$ were in turn used to compute geostrophic velocities that exactly balance these $\hat{\rho}$. On average the correlations between these diagnosed geostrophic velocities and the original input components \hat{u} and \hat{v} were found to steadily decrease from one at the surface, reaching 0.6 at a depth of about 600 m (the latter value is to be compared with correlations of only 0.3 obtained analogously for the diagnosed velocities that would balance the statistically estimated increments $R(T) [\delta\eta^{obs} - \delta\eta^f]$ and $R(S) [\delta\eta^{obs} - \delta\eta^f]$ used in the extended nudging scheme). Even such an approximate balance is found to be sufficient for the re-initialization method to result in a rather smooth model evolution. Moreover, besides having the advantage of sufficiently low computational cost, a local relationship like (5.9) also allows a straightforward application to data sets like tracked altimeter measurements with their intrinsic spatially inhomogeneous data coverage.

Together with the statistically estimated velocity deviations \hat{u} and \hat{v} , the density increment $\hat{\rho}$ provides an approximately balanced and hence dynamically self-consistent estimate for the error of the forecast state $\mathbf{x}^f(t_0)$. The principal idea is to use this error estimate to correct the imperfect model forecast. Obviously, the reliability of the error estimate should be taken into account in the form of some weighting coefficients. For updating the velocity components, as represented by equation (5.1), the weighting coefficients $\mu_0(u_k)$ introduced in the extended nudging scheme, that were just obtained by stressing the analogy to optimal interpolation, should be adopted. In contrast to the

nudging approach, the accuracy of the linear regression scheme for the velocity components here also affects the accuracy of the dynamically derived density increments $\hat{\rho}$. While the relative accuracies of the individual regression coefficients $R(u_k)$ are consistently described by the square of the corresponding correlation, $\text{Cor}^2(u_k, u_{surf})$, the relative error of the vertical derivative (one-sided at the top and bottom of the water column, centered in between), that compares two strongly interrelated quantities, has somewhat arbitrarily been set to the minimum accuracy of the velocity regression coefficients that contribute to the individual derivative. The explicit form for the weighting parameters $\mu_0(\rho)$ thereby becomes

$$\mu_0(\rho_k) = \min_{\substack{u_{up}, u_{low} \\ v_{up}, v_{low}}} (\text{Cor}^2) [1 - r]^2, \quad (5.10)$$

where the subscripts *up* and *low* denote the vertical levels used for computing the individual *z*-derivative, Cor^2 is the corresponding squared correlation with the surface velocity, and *r* is the estimated relative error variance of the altimeter data.

The set of weighted increments $\mu_0(u)\hat{u}$, $\mu_0(v)\hat{v}$, and $\mu_0(\rho)\hat{\rho}$ can then be used to update the forecast state vector \mathbf{x}^f . Although the differences in the weighting coefficients will tend to further degrade the approximate geostrophic balance, the model tolerates a re-initialization based on these increments without problems. Hence it is not necessary to geostrophically balance the weighted increments $\mu_0\hat{u}$, but instead it is sufficient – and physically more sensible – to apply the dynamical balance condition to the physically meaningful quantities \hat{u} (i.e., the model-data misfit projected onto the average superposition of vertical modes). The weighting is only subsequently applied to the self-consistent estimate of the difference between model forecast and observations.

Nevertheless, it remains to partition $\mu_0(\rho)\hat{\rho}$ into increments of temperature and salinity since these, and not density itself, are the state variables of the model. It is evident that for a particular density there is no unique solution to this problem as long as temperature and salinity are treated as independent variables. However, in the ocean these quantities are not really independent, but are closely tied to individual water masses. Once a water mass has been formed and the water is no longer in direct contact with the atmosphere its potential temperature and salinity, and hence potential density, can for many purposes be considered as fixed. Even ocean eddies can to a good approximation be regarded as adiabatic features. In this respect it is only the action of the surface conditions in the formation region that simultaneously determine – and thereby interrelate – temperature and salinity.

The above considerations motivate the use of the conservative aspect of water mass properties to set up the additional constraint required for a unique inversion of $\mu_0(\rho)\hat{\rho}$

for changes of the temperature and salinity forecast fields. Here, it is suggested to require that the assimilation step conserves both temperature and salinity on isopycnal surfaces. This is conceptually rather close to the method introduced by HAINES *et al.* (1993), who assumed conservation of potential vorticity on isopycnals. However, in the CME model the distribution of temperature and salinity on isopycnals within the ventilated thermocline is found to be much more homogeneous than that of potential vorticity. The latter quantity reveals considerable anomalies that are advected by eddies. The different isopycnal structure of these conservative water mass properties can be understood by considering the different setting mechanisms in the water mass formation regions. While temperature and salinity values are essentially fixed by the atmospheric conditions, potential vorticity is also influenced by the oceanic conditions. It is therefore argued that the approach of conserving temperature and salinity on isopycnals is less dependent on an accurately simulated evolution of the ocean prior to the assimilation. However, if only the assimilation period is long enough to cover the ventilation of the thermocline, both alternative conservation schemes are expected to behave very similar.

Conserving temperature and salinity on isopycnals is equivalent to conservation of the local T-S relation of the model forecast. Having to deal with a level model, the algorithm first computes the analyzed density $\rho_k^a = \rho_k^f + \mu_0(\rho_k) \hat{\rho}_k$ at each model level k . Subsequently it searches in the local water column of the forecast state for a water parcel (defined by a T-S point) that, when moved adiabatically to the depth level under consideration, produces just the new density ρ_k^a . In case of a potentially unstable stratification, there may be more than one T-S point that satisfies this criterion, whereupon the algorithm always selects the water that requires the minimum vertical displacement. As long as the density change does not lead to the introduction of new potential densities (with the reference pressure of depth level k) not present in the original water column, this procedure uniquely determines the analyzed temperature and salinity fields. When, on the other hand, the analyzed density exceeds the range of potential densities, no vertically displaced water parcel of the forecast state can acquire the density ρ_k^a . In such cases, that are most likely to occur when light water is introduced at the top, some extrapolation of the original T-S relation towards lower (or greater) densities has to be provided. A simple linear extrapolation of the temperature and salinity differences between adjacent depth levels is not always well behaved when large vertical gradients in temperature and salinity counteract, which turns out to be a particular problem in the seasonal thermocline. For this reason it was decided to increase the vertical step size of the finite difference scheme until the

vertical gradients thus obtained resulted in extrapolated temperature and salinities within some confidence radius of the forecast values (typically 3°C and 0.5 psu).

The re-initialization method introduced in this section may finally be summarized by

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}\hat{\mathbf{H}}(\delta\eta^{obs} - \delta\eta^f) . \quad (5.11)$$

In contrast to the extended nudging scheme (4.11) with its projection matrix \mathbf{H} applied to the surface observations in order to obtain “pseudo” observations of the complete state vector, the operator $\hat{\mathbf{H}}$ of the re-initialization method acts directly on the observed model-data *misfit*. It creates a dynamically self-consistent “pseudo” model-data misfit of the velocity and density fields. As indicated by the results of linear observability theory (section 2.2), we expect just these fields to be observable by altimetry, while additional information is required to determine temperature and salinity individually. It is the function of the operator \mathbf{K} to first weight the estimates according to their accuracy and further invert the weighted density increments $\mu_0(\rho)\hat{\rho}$ for changes in the model state variables temperature and salinity. Because this inversion is based on the conservation of the local T-S relation of the model forecast, the non-linear operator \mathbf{K} strongly depends on the actual forecast state \mathbf{x}^f . For this reason each re-initialization step is computationally more expensive than a single timestep of the extended nudging scheme. However, while a large number of nudging timesteps are required for a smooth model evolution, the approximately balanced analyzed state \mathbf{x}^a can be used to re-initialize the model and hence has to be computed only once per assimilation cycle (e.g., every five days). Because the standard leapfrog time integration scheme would actually require computation of the analyzed state at two successive timesteps, it was decided to restart the model from the analyzed state with an extra Euler backward step, thus reducing the computational cost by a factor of two.

We summarize the description of this second assimilation method with a brief recapitulation of its main conceptual advantages: First, there is no use of rather uncertain statistical correlations between sea surface height and water mass properties. Second, no implicit constraint is imposed on the model climatological mean to remain fixed during the assimilation experiment. Even with the mean sea surface provided by the model, the assimilation of SSH *anomalies* may nevertheless be able to change the *mean* hydrography of the model, for example by controlling intermittent water mass formation. While the first point can be illustrated with a static example, studying possible effects on the model’s mean state obviously requires an extensive assimilation experiment, and shall therefore be deferred to chapter 7.

5.3 Static Performance Tests: A Snapshot

To illustrate the effect of the vertical projection method thus developed, the results of a single update or re-initialization step shall be examined. A map composed of sea surface height anomalies from a full 17 day GEOSAT repeat cycle was produced using the mapping routine described in chapter 6. The period chosen was March 11 – 27, 1987. It is centered on March 19, which on the one hand fits one of the 14 analysis times of the one-year CME model reference run that may be used to provide a typical forecast state \mathbf{x}^f , on the other hand closely matches a hydrographic section from Newfoundland to Bermuda taken by KRAUSS *et al.* (1990) during the period March 23 to April 1, 1987, with a cold core ring near 38°N, 56°W surveyed on March 27 – 28.

In Figure 5.1 *a* the forecast of the sea surface height, produced by the CME model in the pure forward mode without any previous data assimilation, is shown for a small region in the Gulf Stream extension area. Clearly visible is the meandering Gulf Stream with its approximately zonal axis at about 39°N. Also sketched is the position of the Gulf Stream and the cold core ring as estimated from satellite IR pictures (taken from KRAUSS *et al.*, 1990), predicting the Gulf Stream axis about 100 km further north than the model does. Obviously the model forecast does not indicate the presence of a Gulf Stream ring in this region. WILLEBRAND *et al.* (1990) investigated the capability of GEOSAT altimeter data to monitor the surface circulation in this area, and in particular proved detectability of the cold core ring by its associated negative sea surface height anomaly (Figure 5.1 *b*). Hence, it seems appropriate to use this region for testing the assimilation procedure. In this section we shall concentrate on the performance of the instantaneous vertical projection scheme, while the results of the dynamical assimilation experiments with repeated insertion of observations shall be discussed later.

Shown in Figure 5.1 *d* is the difference between the sea surface height anomaly of the model forecast and that observed by GEOSAT. Following the procedure developed in the previous section, this model-data misfit can now straightforwardly be used to compute the so-called analyzed model state \mathbf{x}^a . For a meridional section along 56°W, that according to the findings of KRAUSS *et al.* and also indicated by the GEOSAT data should contain the cold core ring and, of course, the Gulf Stream, the density and temperature fields of both the model forecast as well as of the analyzed state are depicted in Figure 5.2. For comparison also shown are the fields that would emerge from the extended nudging scheme when the width of the nudging time window were reduced to a single timestep.

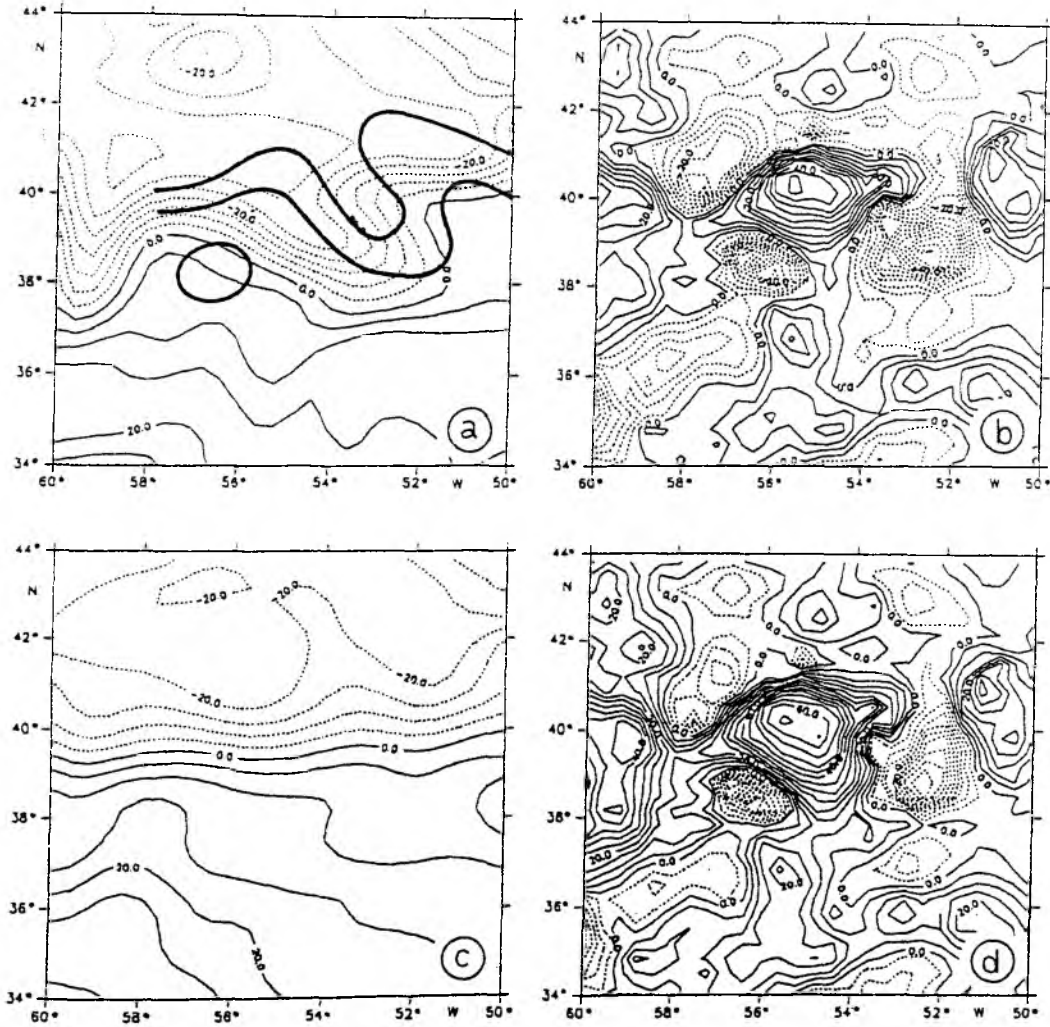


Figure 5.1: (a) Sea surface height, η^f , of the CME model forecast. Also shown is the position of the Gulf Stream and a cold core ring on March 24, 1987 estimated from satellite IR pictures (taken from KRAUSS et al., 1990). (b) Sea surface height anomaly, $\delta\eta^{obs}$, estimated from GEOSAT altimetry data for the 17-day period March 11-27, 1987. (c) Mean sea surface height, $\langle\eta\rangle$, from the three-years model climatology. (d) Model-data misfit $\delta\eta^{obs} - (\eta^f - \langle\eta\rangle)$. The contour interval used is 5 cm throughout. Note the considerable underestimation of mesoscale variability by the model forecast.

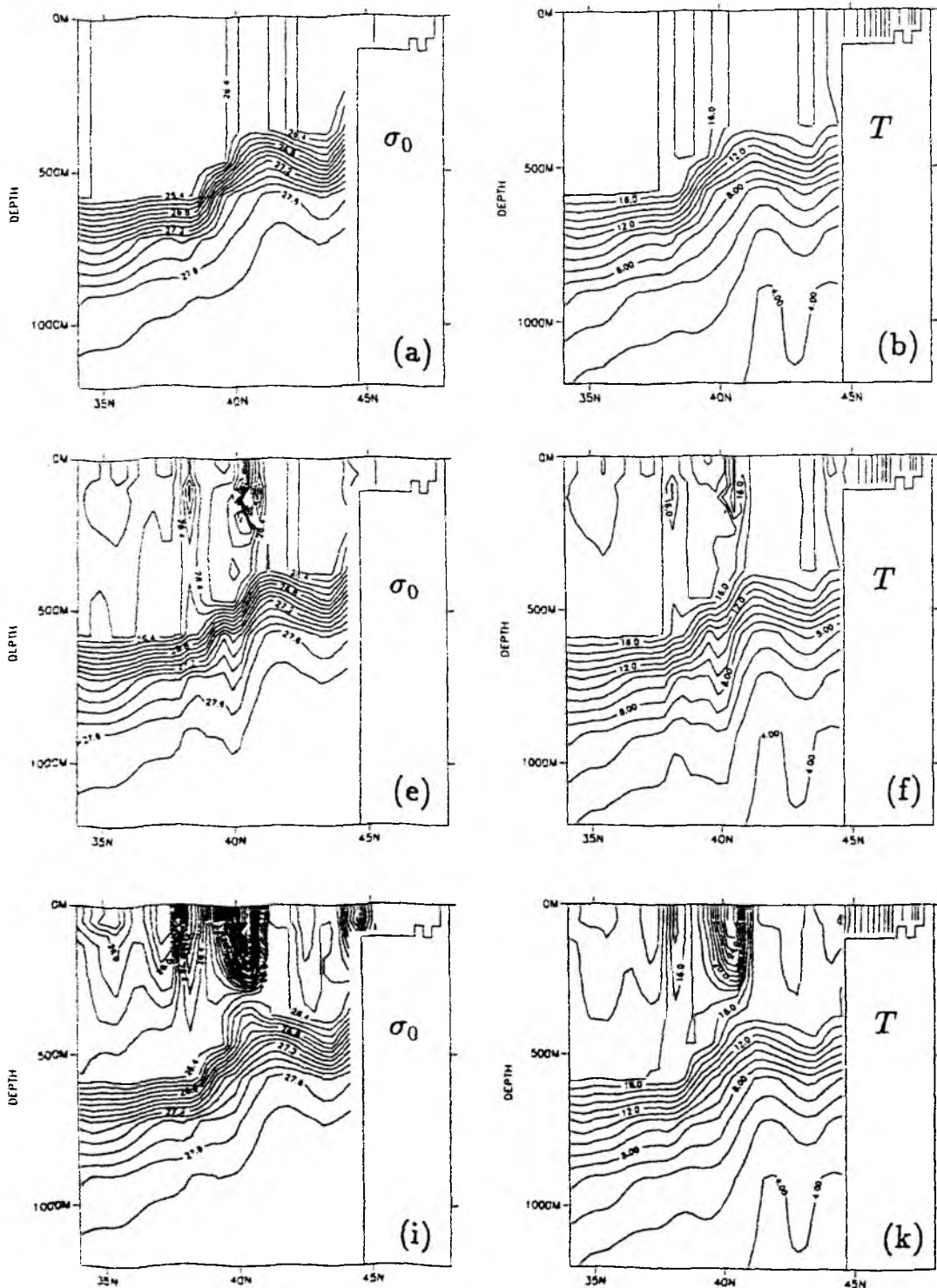


Figure 5.2: Meridional sections along 56°W . Top row: Fields from the forecast x^f for March 19 provided by a CME model run without any previous data assimilation. Middle: Fields from the re-initialization analysis x^a . Bottom: Estimates purely based on the statistical linear regression method. (a),(e),(i) show sections of potential density σ_0 , (b),(f),(k) sections of potential temperature,

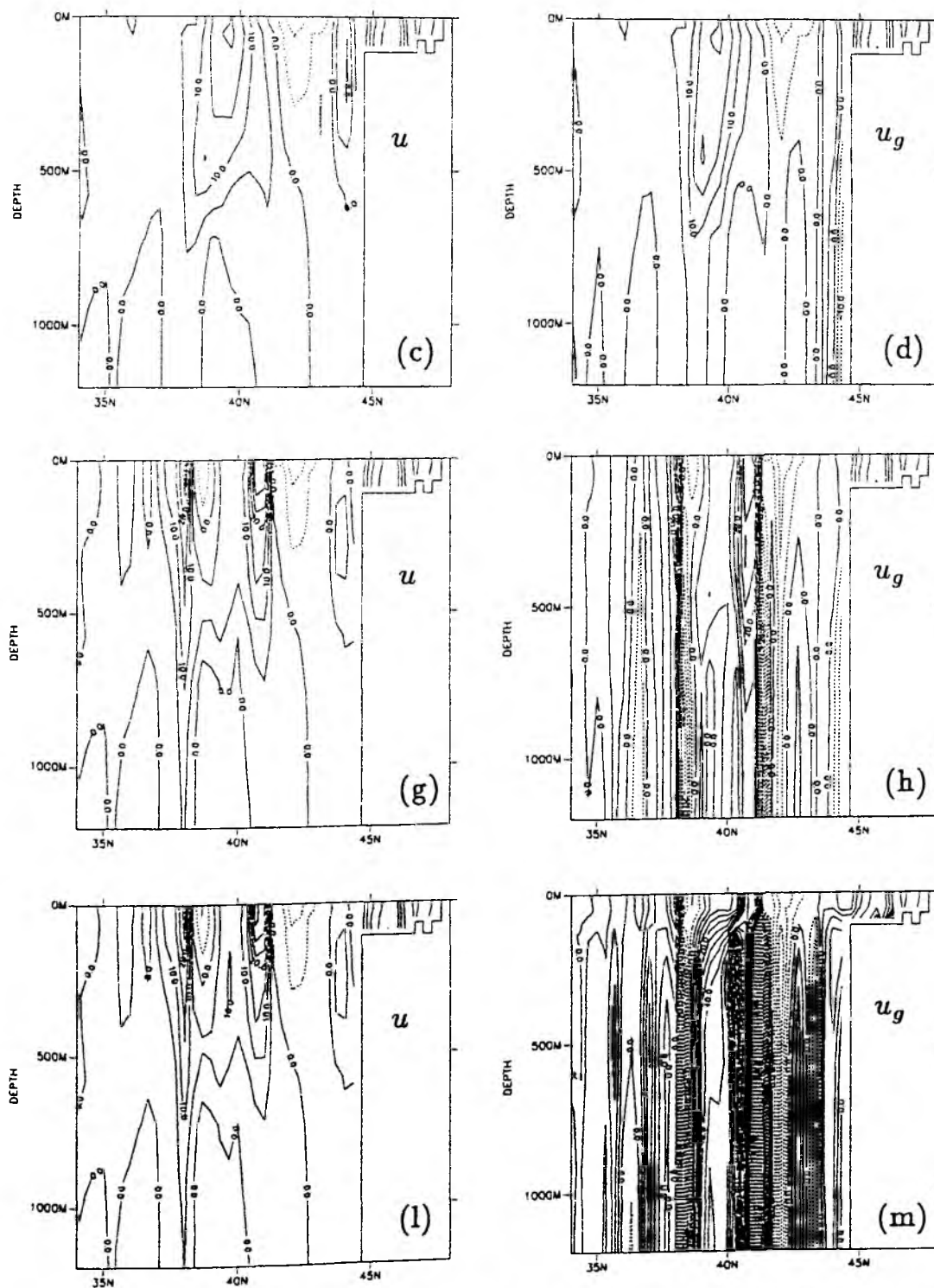


Figure 5.2: (continued): (c),(g),(l) zonal velocity and (d),(h),(m) the diagnosed zonal velocity that would balance the density field of the first column geostrophically, with the absolute surface velocity fixed by the estimates shown in the third column. Contour intervals used are 0.1 for σ_0 , 1°C for temperature and 5 cm s^{-1} for velocity. The GEOSAT data used for the instantaneous updates are shown in Figure 5.1b.

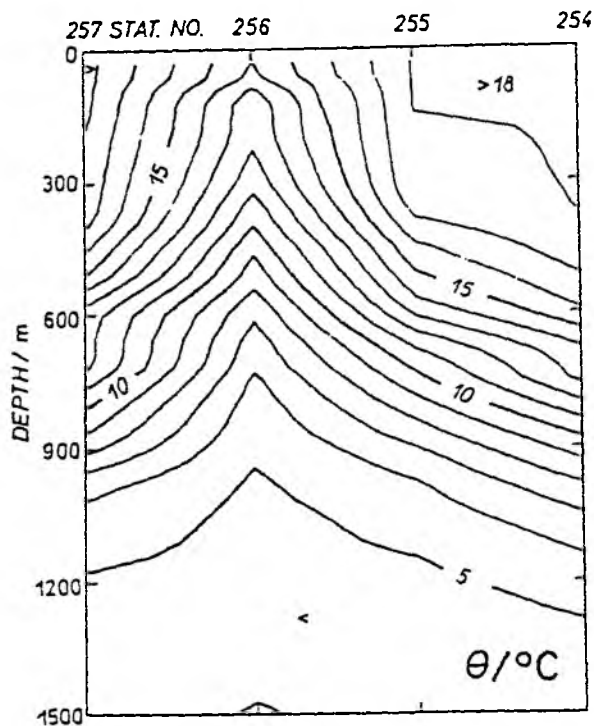


Figure 5.3: *Hydrographic cross section of the cold core ring (reproduced from KRAUSS et al., 1990). The station spacing is about 50 km along a straight line running from southwest (stat. 257 at 37.8°N, 57.1°W) to northeast (stat. 254 at 38.9°N, 55.8°W).*

The forecast fields, presented in the top row of Figure 5.2, show the main front associated with the eastward flowing Gulf Stream between 39°N and 40°N. In accordance with the surface elevation field (Figure 5.1 *a*), there is some indication for a warm core eddy at 43°N, but no evidence for mesoscale variability south of the Gulf Stream. When this forecast state is used as first guess for the assimilation of GEOSAT sea surface height anomalies (Figure 5.1 *b*) following the re-initialization procedure just developed, the analyzed fields shown in the middle row of Figure 5.2 are obtained. The most pronounced changes due to the assimilation are the northward displacement of the strengthened Gulf Stream front, now situated at about 41°N. Also clearly visible is the signal due to the negative sea surface height anomaly at 38°N. While the hydrographic survey taken by KRAUSS *et al.* shows surfacing of the 15°C isotherm in the center of the ring (Figure 5.3), the single-step assimilation does not produce any water colder than 15°C in the upper 100 m. However, the signature of the cold core ring is well developed down to the bottom of the main thermocline. Note that the difference in the density (and hence temperature) fields between forecast and analyzed state is largest not at the surface but at a depth of about 200 m. This is due to the fact that close to the surface there is little vertical shear in the velocity profile given by the regression coefficients $(R(u) + R(v))/2$ used in the inversion of the thermal wind equation (Eq. 5.9). Furthermore, the peculiar density structure at 200 m in the Gulf Stream front can also directly be related to the slight horizontal inhomogeneity

of these regression profiles. However, in the actual dynamical assimilation procedure the convective adjustment scheme of the model code will subsequently be applied to the analyzed fields. This will on the one hand remove any static instabilities present in Figure 5.2 *e*, but might on the other hand also disperse features like the cold core ring. As any convective adjustment of the analyzed fields will affect the model-data misfit of the following assimilation cycle, a repeated assimilation scheme should nevertheless be able to counteract this dispersal by, for example, sequentially introducing more and more cold water while keeping the surrounding water relatively light and warm. We shall investigate the effectiveness of this process for this particular example in chapter 7.

For the moment further concentrating on the static vertical extrapolation of the surface data, we switch to the bottom row of figure 5.2, where the estimated fields $\rho^{est} = \rho(T^{est}, S^{est})$, T^{est} and u^{est} are shown, with

$$T^{est} = T^f + \mu_0(T) [\langle T \rangle + R(T) \delta \eta^{obs} - T^f] \quad (5.12)$$

and

$$u^{est} = u^f + \mu_0(u) [\langle u \rangle + R(u) \delta u_{surf}^{obs} - u^f] \quad (5.13)$$

computed by the linear regression method. These estimates would correspond to the results of a rather hypothetical extended nudging scheme with the Gaussian time window replaced by a delta function centered at the observation time. From the discussion of section 4.3 we recall that for a dynamical application of the extended nudging approach a finite nudging time window is required for a smooth model evolution. This is because, in contrast to the projection scheme of the re-initialization approach, the statistically estimated fields are not properly balanced, as can now easily be understood by comparing the estimated zonal velocities (Figures 5.2 *g, l*) with the geostrophic velocities (Figures 5.2 *h, m*) diagnosed from the estimated density fields shown in the first column. It is evident that the re-initialization estimates shown in the middle row at least approximately satisfy the geostrophic balance condition, while the velocity estimates of the linear regression method show a rather poor resemblance to the geostrophic velocities diagnosed from the statistically estimated density field.

Another striking feature of the fields resulting from the linear regression method is the huge range of temperature and density values close to the surface (Figures 5.2 *i, k*). Temperature estimates exceed 27°C, with corresponding densities of $\sigma_0 < 23.3$. For the region under consideration such water mass properties can certainly not be considered as realistic values. While this deficiency can readily be attributed to the inaccurate

model statistics that considerably underestimate the mesoscale variability, the additional phenomenon that the impact of the linear regression method on subsurface water mass properties quickly decreases with depth is due to the general weak correlations and would also occur in the perfect statistics case.

Although both vertical extrapolation methods show almost identical changes in the zonal currents, it is only the dynamical method of the re-initialization approach that implies concurrent dynamically consistent changes in the density structure down to the main thermocline. We conclude that as far as static assimilation is concerned, the re-initialization approach with its projection technique exploiting the thermal wind equation should be preferred to statistical extrapolation methods like the one used in the extended nudging method. Furthermore, for the assimilation of real altimeter data, the inversion of a statistically estimated velocity field (Fig. 5.2*g*) for a dynamically consistent density field (Fig. 5.2*e*) gives results superior to the opposite conception of geostrophically adjusting the velocity field (Fig. 5.2*m*) to a statistical density estimate (Fig. 5.2*i*). Note that this latter approach is essentially the one proposed by MELLOR and EZER (1991), who obtained quite encouraging results by restricting their investigation to identical twin experiments. Having in mind the rather unrealistic fields shown in Figures 5.2*i*, *m*, this again underlines that one has to be aware of the peculiar tendency of twin experiments yielding overoptimistic results. Nevertheless, because of their unique property of providing objective convergence criteria, such experiments are still important tools for testing assimilation methods. Indeed, we shall continue by investigating the dynamical performance of the above developed alternative assimilation techniques within the framework of identical twin experiments before turning towards the real data case in chapter 7.

5.4 Dynamical Tests Using Identical Twin Experiments

Having investigated the static results of different vertical projection schemes that instantaneously extrapolate the surface observations into the ocean interior, we shall now examine the performance of the actual dynamical assimilation procedure. An appropriate criterion for rating the success of an assimilation experiment is its ability to drive the model state closer towards reality. Since the state of the real ocean is far from being known well enough, we resort to identical twin experiments that enable us to objectively measure the convergence of the assimilation run by computing the difference

between state vectors of the assimilation and control run. Again, for computational economy the simplified box model version described in section 3.2 will be employed. In accordance with the identical twin experiments used in section 4.4 to test different choices for the vertical projection matrix \mathbf{H} of the extended nudging approach (4.11), complete and perfect maps of the sea surface height shall be assimilated every five days. To allow for direct comparison, all assimilation runs use identical initial conditions as well as the same simulated observations. Three basic assimilation runs were performed, with the evolution of the rms difference in the horizontal velocity field at selected levels and in temperature for the whole basin displayed in Figure 5.4. Besides the experiments that employ the extended nudging approach described in chapter four and alternatively the re-initialization developed in the present chapter, for comparison also shown is the result of a model run based on the assimilation technique suggested by MELLOR and EZER (1991).

While the re-initialization method (curves *B*) indicates a slightly better convergence of the surface velocity field than the extended nudging scheme (curves *C*), the temperature field is recovered more accurately by the extended nudging approach. Interestingly, the temperature rms error of the re-initialization method shows a larger drop for the first assimilation cycle (at day 5) than does the extended nudging scheme. However, the slower initial convergence of the latter experiment can be attributed to the large model-data misfit at the beginning of the assimilation experiment and the model's ability to partially adjust to the "pseudo" observations already within the finite nudging time window, thereby reducing the integral effect of the first few assimilation cycles. As already noted in section 4.5, where we compared extended nudging with optimal interpolation spread out in time, this is mainly an initial effect that quickly vanishes as the model state comes closer to the observations. Eventually both the extended nudging approach and the re-initialization method produce almost similarly encouraging results.

Also shown in Figure 5.4 are the results of an experiment using the assimilation method proposed by MELLOR and EZER (1991). Temperature and salinity changes are derived by the statistical linear regression scheme, with the changes in the velocity field being adjusted geostrophically to the weighted density increments. However, to ensure an exact obedience of the balance condition, no weighting factors are applied to the geostrophic velocity increments. This is in contrast to the re-initialization method developed in this chapter, where it was decided to downweight the geostrophically adjusted density increments $\hat{\rho}$ by their estimated accuracy $\mu_0(\rho)$. Despite the fact that these weighted increments are not exactly balanced, the model tolerates a re-

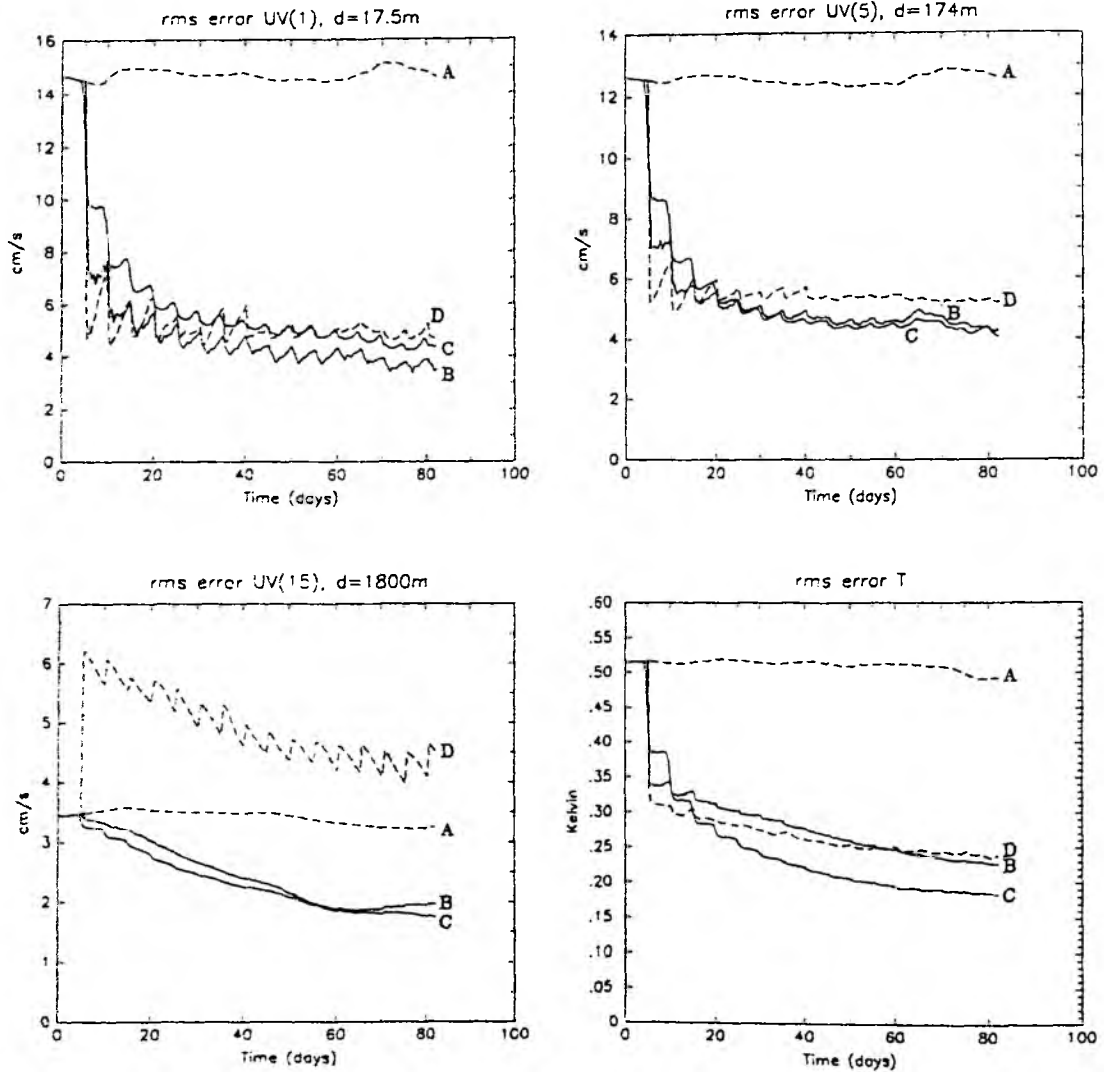


Figure 5.4: Time evolution of the rms difference between control run and different assimilation runs. Values for velocities are shown for selected depth levels (17.5 m, 174 m, 1800 m), while the rms difference for temperature is computed over the whole basin. Curves (A) denote the reference run where no data are assimilated, (B) to the re-initialization method, (C) refer to the extended nudging scheme, and (D) to the approach suggested by MELLOR and EZER (1991).

initialization (although some inertia-gravity waves are excited as can be deduced from the somewhat noisy evolution of the rms curves *B*). The results of these different re-initialization methods show considerable differences, which are most pronounced in the deep velocity field, where application of the MELLOR and EZER method even increases the rms error! This rather unpleasant phenomenon can be traced back to the fact that any errors in the statistically estimated density profile (due to imperfect correlations) create errors in the vertical shear of the geostrophic velocities. When integrating this shear profile beginning from the directly observed surface velocities, this error accumulates, possibly resulting in completely wrong estimates for the subsurface currents even in the perfect statistics case.

It is interesting to note that except for the deep velocities, the initial reduction of the rms error is largest for the assimilation method of MELLOR and EZER. Nevertheless, the poor estimates of the deep velocity field have a dynamical effect that leads to a sharp increase in the rms error of the surface velocity field between successive re-initializations. This is in contrast to the more cautious approach of using estimated increments that are downweighted according to their reliability, as suggested in this study. It allows the model to sequentially accumulate strictly the reliable part of the extrapolated model-data misfit. It also emphasizes the model's dynamical role to absorb, store and interpolate information contained in a time series of observations. The efficiency of both the extended nudging approach and also the re-initialization method just developed is illustrated by the low rms error level at the end of the assimilation experiments.

While the identical twin experiments with their inherently perfect statistics do not indicate a clear distinction between the two methods proposed in this study, they already show a significant improvement over the assimilation technique of MELLOR and EZER. However, recalling the results of the static example presented in the previous section, we expect pronounced differences between the extended nudging and the re-initialization method when GEOSAT data are combined with statistical correlations deduced from the CME model with its still systematic deficiencies.

Chapter 6

The GEOSAT Data Set

The U.S. Navy's altimetric satellite GEOSAT was launched in March 1985 with the primary goal of accurately mapping the global marine geoid for so-called defence purposes. This objective was accomplished along a dense grid of non-repeating ground tracks during the first 18 months of operation. After completion of this classified Geodetic Mission, the satellite was maneuvered into an exact repeat orbit with ground tracks nearly coincident with those of NASA's civilian satellite SEASAT. Consequently, the – for this very reason – unclassified data of this Exact Repeat Mission (ERM) do not contain any new geoid information. The ERM began on November 8, 1986, and ended on September 20, 1989, with complete failure of all on-board data storing systems. With an inclination of the orbital plane of 108° , the ERM covers the latitude range from 72°S to 72°N . The satellite's altitude is 800 km and each revolution takes 100.6 min. There is a total number of 244 ground tracks, which are repeated within ± 1 km every 17.05 days. Successive ascending nodes are displaced by about 25° to the west, resulting in an eastward moving three-days subcycle with adjacent tracks being sampled three days apart. The final grid spacing is 164 km at the equator and about 120 km at mid latitudes. The altimeter measurements taken during GEOSAT's ERM are distributed as 1-s averages (corresponding to an along-track sampling distance of 6.8 km) by the NODC (National Oceanographic Data Center) in the Geophysical Data Records (CHENEY *et al.*, 1987) that also contain internal information about data quality as well as various external correction estimates, e.g. for tides and atmospheric effects.

An example of the valid 1 s^{-1} measurements over the Atlantic Ocean is shown in Figure 6.1 for one 17-day repeat cycle. Note the relatively large data gaps for a number of tracks. This quite significant data loss is to a large part ascribed to the use of a passive gravity-gradient stabilization to control the attitude of the spacecraft by

connecting the satellite bus via a long boom to a counter-mass (CHELTON *et al.*, 1990). Due to different geometrical cross sections of bus and counter-mass, solar radiation pressure induces a small torque on the satellite. This may eventually lead to attitude excursions exceeding the tolerable range of $\pm 1^\circ$, fixed by the antenna beam width of 2° . With increasing sunspot activity of the solar cycle, the data loss increased from about 5% at the beginning of the ERM to a level near 60% in the fall of 1989.

6.1 Data Preprocessing

With the results of the identical twin experiments indicating that the convergence timescale of the assimilation is typically of the order of a few weeks, it was decided to limit the computationally expensive assimilation of GEOSAT data into the CME model of the North Atlantic Ocean to a period of one year. Because loss of altimeter data became more frequent towards the end of the GEOSAT mission, the year 1987 was chosen as assimilation period. Actually, the data set used in this study covers the entire North Atlantic Ocean between 10°S and 65°N from November 1986 to January 1988. It consists of data preprocessed by STAMMER (1992), who first applied various corrections using the estimates that were provided with the Geophysical Data Records. These included corrections for solid Earth and ocean tides and also for ionospheric and atmospheric path delay using the dry and wet tropospheric corrections derived from the Fleet Numerical Oceanographic Center model. Further corrections were applied for the inverse barometric effect and also for the sea state bias using 2% of the significant wave height (SWH). To eliminate errors due to local effects like sea ice or rain cells, any data that satisfied one of the following criteria were excluded: (1) SWH exceeding 10 m, (2) standard deviation of SWH (within the 1 s interval) exceeding 10 cm, (3) standard deviation of the automatic gain control exceeding 0.2 dB, or (4) standard deviation of SSH exceeding 10 cm. Since tidal corrections for continental shelf regions are somewhat uncertain, finally any data for ocean depths less than 1000 m were excluded.

After these corrections have been applied, the altimeter data are still dominated by the geoid signal with typical amplitudes of 50 m and by the rather large radial orbit error (4 m rms) of the NAG GEM-10 orbit (CHENEY *et al.*, 1987). To extract the time varying ocean signal for the entire region, the collinear analysis technique (CHENEY *et al.*, 1983) was applied separately in subareas, each extending over 40° - 50° in latitude, with adjacent areas overlapping by 10° in order to avoid boundary effects. With some care taken to avoid a bias in the mean estimate due to data gaps, the

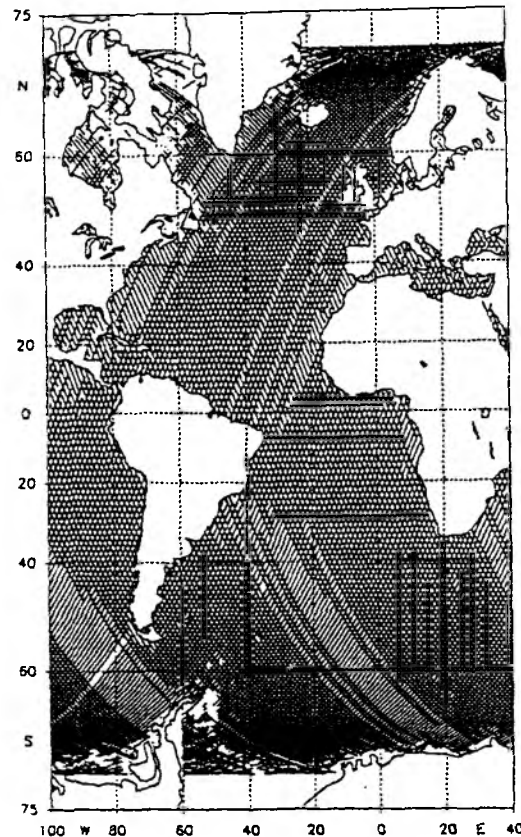


Figure 6.1: *Ground tracks of all valid 1 s^{-1} GEOSAT measurements during the 8th repeat cycle from March 7 to March 23, 1987.*

ensemble mean of the first 43 repeat cycles (i.e., the first two years of the ERM) was subtracted from the individual tracks, thus eliminating the geoid as well as the signal due to the steady ocean circulation. To eliminate the orbit error – and inevitably also some oceanic signal at long wavelengths – the least square polynomial over arc lengths of typically 5000 km was removed from the residual signal of each individual repeat. (In the case of arcs shorter than 1000 km only bias and trend were removed instead.) Finally, an along-track filter was applied and the data were subsampled every 20 km. The above corrections and their uncertainties are discussed in detail by STAMMER (1992), who also provided these corrected GEOSAT altimeter data for the present study.

Because the vertical projection schemes, that turned out to be necessary for a reasonable convergence of the assimilation twin experiments (see chapter 4.4), are defined at the individual columns of grid points, some interpolation of the irregularly distributed satellite altimeter data onto the model grid is appropriate. In order to produce a

time series of two-dimensional quasi-synoptic maps, a linear space-time objective analysis technique (BRETHERTON *et al.*, 1976) was employed. Note that by choosing a small enough time window for the mapping procedure, one still retains the option of (quasi) along-track assimilation. To reduce the computational burden of interpolating $O(10^5)$ GEOSAT observations onto some 5×10^4 model surface grid points, a suboptimal method had to be chosen. The algorithm used is based on an improved version (NADA 5, provided by P. DEMEY) of the original version described by DEMEY and MÉNARD (1989). It greatly reduces the number of data points for the objective analysis by considering only statistically independent data within a chosen influence radius (typically 400 km and 30 days). Like any objective analysis algorithm it nevertheless requires some *a priori* information in the form of a correlation function.

STAMMER and BÖNING (1992) showed that the spatial autocorrelation functions, averaged over 10° by 10° boxes, of the GEOSAT sea surface height anomalies are not significantly different for ascending and descending tracks. Because of the east-west symmetry of the track pattern, this fact alone does not exclude the possibility of significant differences in the zonal and meridional length scales. However, they also found generally good agreement between these autocorrelation functions and the analytic functions estimated from the local POLYMODE (DEMEY and ROBINSON, 1987) and TOURBILLON (ARHAN and COLIN DE VERDIÈRE, 1985) experiments (Figure 6.2a). Based on this evidence, in the present study the same analytic isotropic form of the spatial autocorrelation function as for the POLYMODE experiment was used for the mapping:

$$C(r) = \left[1 + ar - \frac{1}{3}(ar)^3 \right] e^{-ar} \quad (6.1)$$

where $a = 2.1038/L_0$, and L_0 is the lag of the first zero crossing. As further shown by STAMMER and BÖNING, L_0 decreases rather smoothly from 255 km at the equator to 90 km at 60°N with only little zonal variations. An analytic model for the lag of the first zero crossing L_0 was therefore fitted to the zonal averages of the GEOSAT values (see Figure 6.2b):

$$L_0(\lambda, \varphi) \equiv L_0(\varphi) = 50 \text{ km} + 205 \text{ km} \cdot \frac{\varphi_0^2}{\varphi^2 + \varphi_0^2} \quad (6.2)$$

with $\varphi_0 = 30^\circ$. The spatial autocorrelation function modelled by the above equations is then multiplied by a stationary Gaussian decay, $D(t) = \exp(t^2/T^2)$. If only a single SSH map were to be assimilated, the e-folding decay time T should in principle also be derived from the GEOSAT data set. However, here we will assimilate a time series of such maps. To avoid assimilation of the same information again and again (and

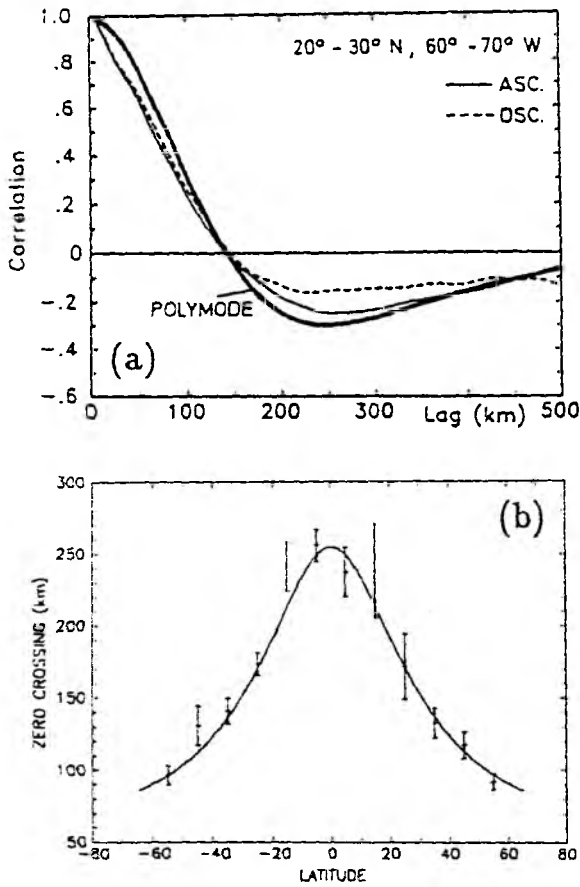


Figure 6.2: (a) GEOSAT mean autocorrelation function from ascending (solid line) and descending (dashed line) tracks in the 10° by 10° box 20° - 30°N, 60° - 70°W. Included is the analytic autocorrelation function (6.1) representing typical POLYMODE sea surface stream function statistics (DEMEY and ROBINSON, 1987). (Figure taken from STAMMER and BÖNING, 1992.) (b) Zonal averages of the lag of the first zero crossing of the GEOSAT SSH autocorrelation function (from STAMMER and BÖNING). Also shown is the analytic fit $L_0(\varphi)$ used for the mapping routine.

thereby violate the statistical assumptions underlying the construction of the weighting coefficients), the decay time T should be related to the time interval ΔT between successive maps. We shall somewhat arbitrarily set $T \approx 2\Delta T$ and investigate the effect of different choices for the time interval ΔT subsequently. With regard to an instrumental noise level of about 3.5 cm rms (SAILOR and LESCHACK, 1987), the GEOSAT data exhibit a relatively low signal to noise ratio for large parts of the North Atlantic Ocean (e.g., STAMMER *et al.*, 1991). We therefore finally assume a white noise level of 30% of the local process variance (within the chosen influence radius). The output of the objective analysis package consists of an estimated SSH field of the entire model domain, and a map of the estimated relative error variance $r = \sigma^2(\eta^{obs})/\text{Var}(\eta^{obs})$, that enters the weighting coefficients of both the extended nudging approach (chapter 4) and the re-initialization method (chapter 5).

6.2 Maps versus Tracks

The basic question to be investigated in this section is whether a “real time” assimilation of the incoming tracks or a less frequent assimilation of analyzed maps combining all tracks within a certain time window should be preferred. While only the former approach retains the full temporal resolution of the data and leaves the horizontal interpolation to the model dynamics, the latter one favours statistical interpolation to produce a map with a much higher spatial resolution, that can then be consistently provided to the model at once.

KINDLE (1986) first examined this question using a simple updating scheme for a one-layer reduced gravity model. For simulated SSH measurements along a hypothetical 72-day repeat orbit he combined tracks within different time windows of either 3 days, 6 days, 12 days, or 18 days for updating steps at the same intervals. Reduction of the rms error was largest for the 12 days assimilation cycles. More frequent updates lead to a buildup of spurious gravity noise, reflecting the fact that the employed update scheme was not well adapted to the assimilation of unbalanced data. Using an update scheme for the surface-layer streamfunction of a quasi-geostrophic model, BERRY and MARSHALL (1989) also found that along-track assimilation of data from a simulated 14-day repeat orbit was not sufficient to constrain the model. In contrast, HOLLAND and MALANOTTE-RIZZOLI (1989) showed for a number of orbit schedules that nudging of the along-track vorticity component enabled their quasi-geostrophic model to reconstruct the ocean state quite accurately. WHITE *et al.* (1990 a,b,c) also successfully assimilated tracked altimeter data into quasi-geostrophic models by employing a statistical interpolation scheme that, however, accounted for some non-local error covariance within a narrow swath containing the respective ground tracks.

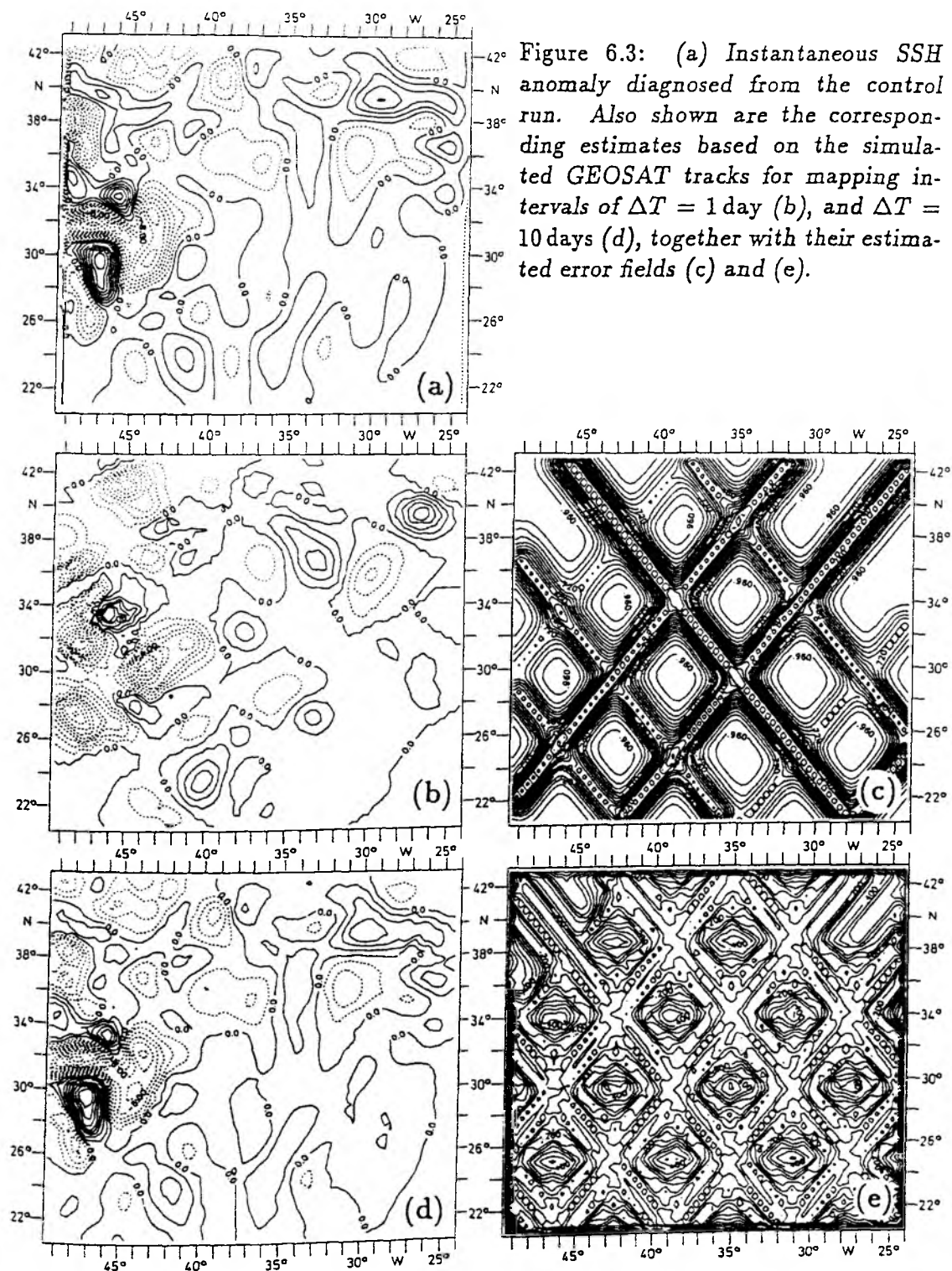
The emergent picture is that, for the various suboptimal assimilation methods used, a general rule whether maps or tracks should preferably be provided to the assimilation scheme does not exist. Note however, that except for the early work by KINDLE (1986) using a rather crude assimilation method and a simple ocean model, none of the above studies has thoroughly investigated the maps-versus-tracks question for a given tracked data set. In order to utilize the GEOSAT data set correctly, we shall therefore try to examine this question with regard to the assimilation schemes developed in the previous chapters.

With the information content of the data set being fixed, it remains to determine an appropriate way to provide this information to the model evolving in time. Actually,

the mapping procedure described above includes additional *a priori* knowledge in form of the isotropic correlation function which would be neglected by a pure along-track assimilation. For reasons of straight comparability it is but appropriate to ensure that assimilation of along-track observations and of preprocessed maps both contain the same amount of information. It was therefore decided to apply the mapping routine to the individual tracks as well. This can be reached by decreasing the time interval ΔT between successive maps (together with the *e*-folding decay time $T = 2\Delta T$). In this way the problem is reduced to examining different choices of ΔT .

To investigate the optimal range of ΔT , a series of identical twin experiments was performed. The model used was the simple four-level box model configuration of chapter 3.3. First, a synthetic satellite orbit was implemented that approximated the sampling characteristics of the GEOSAT ERM. To avoid interpolation of the simulated model SSH on the satellite ground track, the tracks were chosen to have a slope of one on the model grid, corresponding to a cross-over angle of approximately 72° at 30°N . The 244 tracks are separated by four grid points in east-west direction (roughly 130 km). Successive tracks are displaced by 68 grid points to the west, and a full repeat cycle takes 17 days. Each pass of the satellite over the model domain is assumed to happen instantaneously at the model timestep ($\Delta t = 0.5\text{ h}$) closest to the actual passing time. SSH anomalies (deviations from the model climatology) were stored for all grid points along the track, corresponding to an along-track sampling distance of about 55 km.

The simulated satellite altimeter was then flown over the control model ocean, yielding a tracked data set that further served as input for the objective analysis package described above. Five different sets of analyzed maps were produced, with corresponding mapping periods ΔT of 1 day, 2 days, 5 days, 10 days, and 15 days, respectively. Setting the *e*-folding decay time T equal to $T = 2\Delta T$ ensured that the total information content of any timeseries of maps is about the same. No mapping intervals shorter than $\Delta T = 1\text{ day}$ have been investigated, but as indicated by the estimated error map of Figure 6.3c, the choice $\Delta T = 1\text{ day}$ can be considered to be quite a good approximation to direct along-track assimilation. It is evident from Figure 6.3b that the individual maps obtained with a short mapping interval and the associated short *e*-folding decorrelation time miss important features like the positive anomaly at 29°N , 47°W of the “true” SSH anomaly (Figure 6.3a). Nevertheless, since the nudging coefficients (Eq. 4.34, 4.35) take into account the estimated error variance r , only the reliable parts of the respective map are assimilated into the model. Increasing the mapping interval and hence the decorrelation timescale leads to more homogeneous error maps (Figure 6.3e) along with more complete estimates of the SSH anomaly



field (Figure 6.3 d) that, however, are provided to the model less frequently.

As shown previously in the context of identical twin experiments both the extended nudging and the re-initialization method give about the same results (section 5.4). Because it is easier to implement, the extended nudging approach was chosen for this series of twin experiments. Thus, all experiments used exactly the same assimilation method, but different mapping and hence assimilation time intervals ΔT . The evolution of the rms difference between assimilation runs and control run is presented in Figure 6.4. Obviously, the reduction in the rms error is largest when mapping intervals of 5 to 10 days are used. Of all the experiments performed, it is the $\Delta T = 1$ day run, simulating along-track assimilation, that clearly exhibits the worst convergence! It appears that the model can not adequately convert asynoptic observations into an accurate synoptic hindcast state. On the other hand increasing the mapping interval ΔT beyond 10 days again deteriorates the convergence of the assimilation run towards the control model. This is not too surprising, as $2\Delta T = 20$ days is about the e -folding decorrelation time of the SSH field of the model – and also of the GEOSAT data (e.g., LETRAON, 1991). Quite generally, information is lost when the objective analysis overestimates correlations between distinct observations.

Interestingly, the differences between the above experiments are most pronounced in the convergence of the velocity fields. This may be understood by considering that the velocity components are related to the gradient of the mapped SSH field, and hence depend much more on non-local information than temperature estimates, which the linear regression scheme relates to the local sea surface height itself. While an along-track assimilation scheme determines the cross-track velocity components very accurately, any other velocity estimates are rather poor. Note that applying the mapping routine to each track individually implies the assumption of vanishing along-track velocity. Hence, the kinetic energy of the surface current field will be underestimated. It actually turns out that after 90 days of assimilation the basin averaged kinetic energy has decreased by 40% for the $\Delta T = 1$ day experiment, compared with 25% for the $\Delta T = 5$ days experiment! (In all the extended nudging assimilations performed, a general decrease of kinetic energy is observed, being due to two reasons: First, the nudging terms always acts as a damping term, second, the extended nudging scheme nudges towards *mean* anomalies that explain only the fraction Cor^2 of the variance. See also chapter 7.2.) We therefore argue that it is mainly the large error in the estimated velocity field which is responsible for the low efficiency of along-track assimilation. Because of its even greater dependence on accurate velocity estimates, this result is expected to hold for the re-initialization method as well.

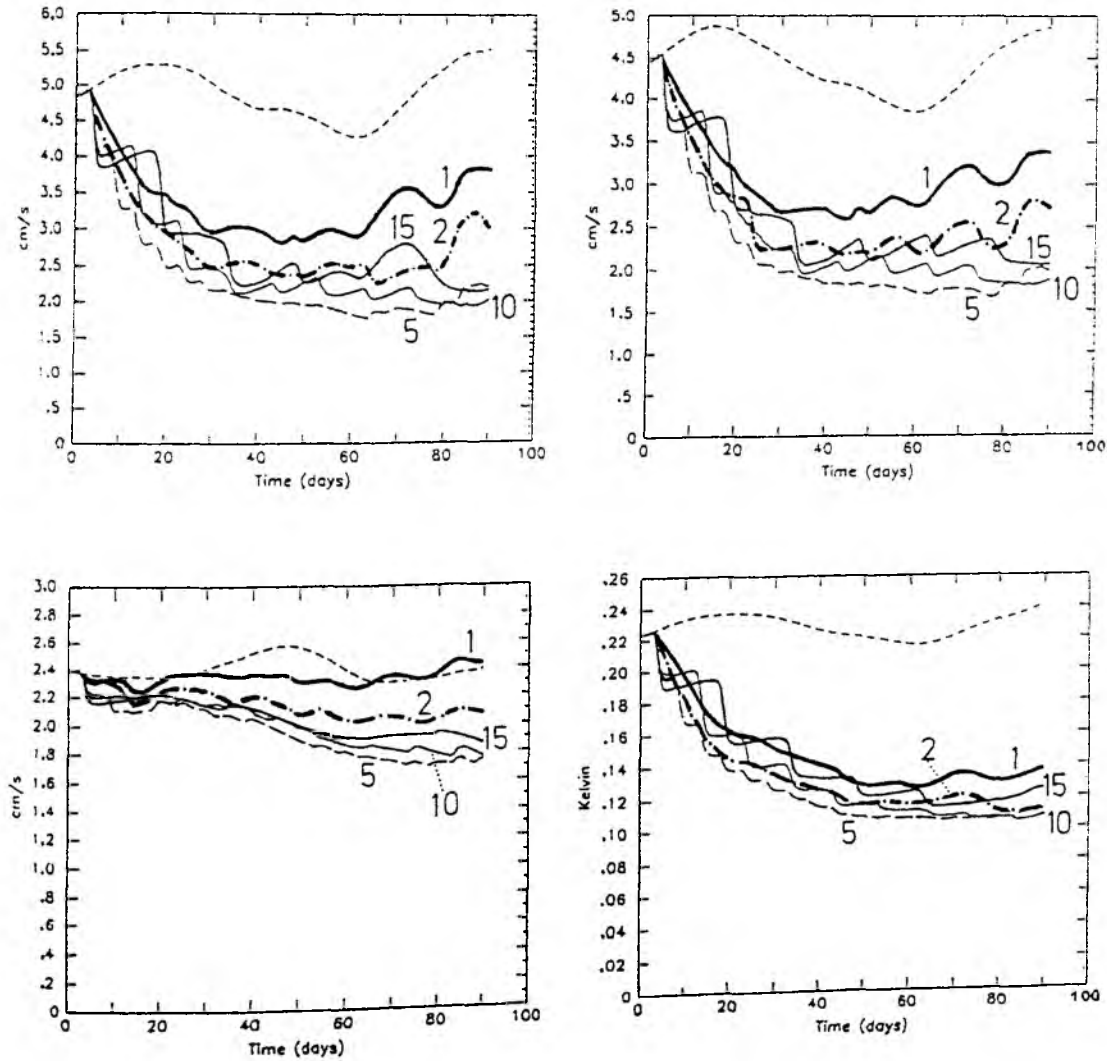


Figure 6.4: Time evolution of the rms differences between control run and assimilation run. Velocity values are shown for selected depth levels and the temperature rms difference is computed over the entire basin. Different mapping (and hence assimilation) cycle lengths ΔT are indicated by the curve labels (ΔT in days). The dashed curve refers to a run where no data are assimilated.

From the above experiments we deduce that, when applying our simple assimilation methods to combine altimeter data and primitive equation models with prognostic momentum equations, the use of preprocessed maps of SSH should be preferred to a direct along-track assimilation. Apparently, the suboptimal sequential assimilation methods developed in this study work best when prior to the actual assimilation a rather complete estimate of the entire model state vector is provided. This is presently reached by separating vertical correlations from horizontal and temporal ones: First, the mapping routine exploits horizontal and temporal correlations independently of the model state. The estimated surface field is then extrapolated vertically, either by the pure statistical projection scheme of the extended nudging approach, or by the dynamically more consistent inversion of the thermal wind equation used for the re-initialization method. The time period between successive maps should be chosen large enough to allow for a rather complete coverage in space, but short when compared to typical time scales of the process. For the GEOSAT data set the results obtained by LETRAON (1991) indicate e-folding decorrelation times of roughly 20 days. Combining this estimate with the above displayed twin experiment results, it was decided to use a mapping interval of $\Delta T = 5$ days together with an e-folding correlation decay time of 10 days to produce the sequence of GEOSAT SSH anomaly maps for the year 1987. An example of an individual map is given in Figure 6.5, that also shows the corresponding map of the estimated relative error $r = \sigma^2(\eta^{obs}) / \text{Var}(\eta^{obs})$.

We finally note that the recommendation of using analyzed maps instead of assimilating directly along the incoming tracks, depends on the assimilation scheme as well as on the circulation model chosen. Both suboptimal assimilation methods developed in this study extrapolate the surface observations into the ocean interior for each column of grid points separately. Neither horizontal nor temporal correlations are used in this static estimation process. However, in the ocean such correlations clearly exist, and any optimal assimilation scheme takes these into account, either dynamically (adjoint method) or statistically in the form of non-local error covariances (Kalman filter). Hence, for optimal methods a direct assimilation of satellite altimeter data along-track may nevertheless be appropriate (e.g., SCHRÖTER *et al.*, 1993).

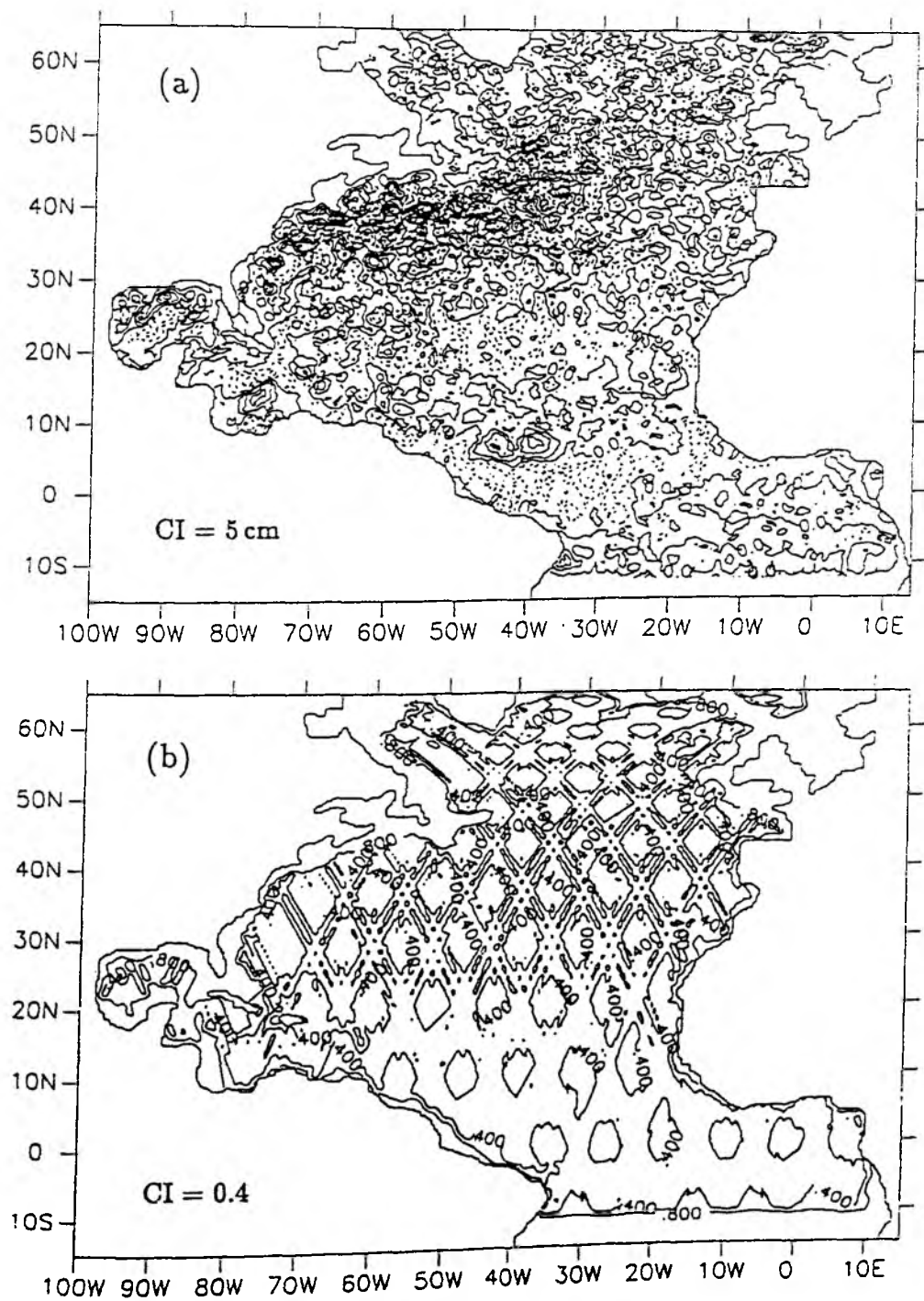


Figure 6.5: (a) Objectively analyzed GEOSAT SSH anomaly field for March 19, 1987, and (b) corresponding estimated relative error variance.

Chapter 7

Results of the Assimilation Experiments

The two assimilation techniques developed in chapters four and five, respectively, shall now be used to combine the GEOSAT data record with the WOCE CME model of the North Atlantic Ocean. A sequence of objectively analyzed maps of satellite measured SSH anomalies, covering the year 1987 at 5-day intervals, is assimilated into two one-year runs of the eddy-resolving CME model. The first assimilation run employs the extended nudging approach of chapter 4, while the second one uses the re-initialization method developed in chapter 5. Both numerical experiments begin with identical though wrong initial conditions on January 1, 1987. Actually, this initial state is the restart state of the model year 32, which can therefore serve as the one-year reference run with no data assimilated. To allow for straight intercomparison between this reference run and the two assimilation experiments, the same forcing is applied throughout. As mentioned in section 3.1, the forcing fields are derived from climatology, and hence will not be fully consistent with the GEOSAT observations for the year 1987. In particular, the wind forcing is interpolated linearly between monthly mean values, omitting any higher frequency forcing. However, experiments with daily forcing (based on data from the European Center for Medium Range Forecasting) indicate that high frequency wind components cause a strong inertial response in the surface layers of the model, but do not significantly affect the eddy variability of the ocean interior (BÖNING, 1992). This supports the supposition that in the North Atlantic Ocean it is mainly the action of instability processes along frontal zones which generates mesoscale variability. In this respect it can already be anticipated that in order to improve the model's representation of eddy variability, the hydrographic situation has to be properly accounted for by the assimilation procedure. Slight inconsistencies in the actual wind

forcing may on the other hand be regarded as acceptable for our purposes of exploring feasible assimilation methods.

To avoid problems that can arise from applying the geostrophic assumption, inherent in the assimilation schemes developed so far, close to the equator, no data are assimilated south of 10°N . This is not considered to be a serious restriction, since we expect the largest impact of the assimilation of altimeter data in the subtropical and subpolar regions of the Atlantic Ocean. The differences between the results of the CME model with no data assimilated and GEOSAT observations are most pronounced in these regions (STAMMER and BÖNING, 1992), whereas for the equatorial region DIDDEN and SCHOTT (1992) already found a rather close agreement between model results and GEOSAT data.

In this chapter I will present the main results of the two assimilation experiments. Advantages and deficiencies of both assimilation methods will be revealed, beginning with the surface height field and the distribution of eddy kinetic energy. Quite important conclusions will then emerge from the analysis of the very different effects the two assimilations methods have on the model hydrography. Finally, it will be shown that the re-initialization method developed in chapter five of this study does indeed lead to promising changes in the *mean* state of the model by assimilating temporal *anomalies* of the surface elevation.

7.1 Sea Surface Height

First of all, we will analyze the effect of the assimilation of GEOSAT data on the surface height field. As mentioned in the third chapter, the CME model has a rigid lid. Consequently, surface pressure, or equivalently surface height, is not a prognostic variable, but has to be computed diagnostically from the model state. Since the diagnostic derivation essentially involves the full state vector, adjustment of the SSH field of the assimilation model towards the observations is not trivial, but actually requires some dynamical consistency of the assimilation procedure. In the present application both assimilation experiments use the model mean surface to reference the GEOSAT SSH anomalies. While the re-initialization method could in principle use any given mean surface instead, the extended nudging approach more restrictively requires a complete mean state of the model consistent with the chosen mean surface. In any case, not being aware of any accurate, high-resolution mean surface height estimate for the entire model region, we just retain the mean sea surface simulated by the

CME model.

A first criterion of the assimilation experiments' performance can be obtained in close analogy to the standard analysis of identical twin experiments: We simply compute the rms difference between the SSH maps derived from the GEOSAT data set and those diagnosed from the individual assimilation experiments. Surface height fields are computed from a set of 14 model states, separated by time intervals of 26 days. A corresponding set of GEOSAT maps is obtained by treating the data from a complete 17-day repeat cycle centered at the respective analysis time as synoptic. All grid points north of 10°N (i.e., within the assimilation region) with an estimated relative GEOSAT error variance $r < 0.5$ are considered. The rms differences between the GEOSAT maps and the three CME model experiments (two assimilation runs, one reference run with no data assimilated) are shown in Figure 7.1:

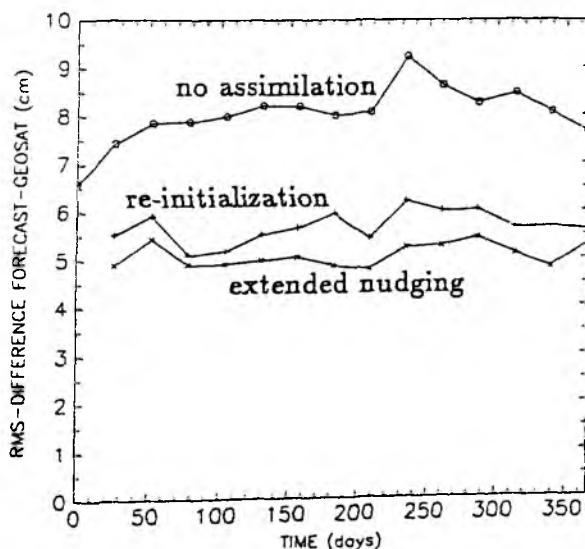


Figure 7.1: *Temporal evolution of the RMS difference between objectively analyzed GEOSAT maps and the different numerical model experiments, respectively. Only grid points north of 10°N with estimated error variance $r < 0.5$ are taken into account.*

The original rms difference between GEOSAT maps and the CME model with no data assimilated amounts to roughly 8 cm. Both assimilation methods can reduce this difference by about 35%, with somewhat better convergence of the extended nudging approach (remaining difference of about 5 cm rms) when compared with the re-initialization method (5.5 cm rms). Consistently with the rapid convergence of the surface fields in the identical twin experiments, these error levels have already been reached at the first analysis time (day 26) and remain quite stable during the one-year

experiments. The seeming advantage of the extended nudging scheme is even more pronounced in the explained variance: On average, the extended nudging experiment can explain 26% of the variance in the GEOSAT maps, while the re-initialization experiment does only account for 17%. Following a rather common tradition and using these statistical figures to rate the performance of assimilation procedures, there would be little hope for the re-initialization method to survive a competition with the extended nudging scheme. Nevertheless, it will subsequently become clear how completely misleading such a diagnosis through statistical measures can be, if one is eventually interested in physical processes in the ocean.

Figure 7.2 displays the rms variability of the surface height field, computed from the individual one-year experiments. The rms surface variability of the CME model with no data assimilated (Fig. 7.2a) shows maximum values of about 22 cm near 35°N, 70°W, where the separating Gulf Stream forms a large, unrealistic loop before turning eastward. The region of maximum variability then follows the path of the Gulf Stream and the North Atlantic Current (NAC), that does not properly cross the Mid-Atlantic Ridge (MAR), but instead forms a spurious band of high variability in the Irminger Sea southeast of Cape Farewell. The enhanced variability in the tropical Atlantic is associated with the seasonal cycle of the North Equatorial Counter Current (NECC) and the retroflecting North Brazil Current, that turns toward the NECC during summer. As shown by SCHOTT and BÖNING (1991), there is good correspondence between model results and observational evidence in this region.

It is mainly the region north of 30°N, where the differences in the spatial distribution of the rms surface variability between the original CME model run and the corresponding GEOSAT maps are most pronounced. The rms variability inferred from the GEOSAT data (Fig. 7.2b) is largest in the Gulf Stream extension along 40°N between 70°W and 40°W. Further downstream there are two branches of high variability crossing the MAR, one corresponding to the NAC at about 50°N, while the other one at 35°N is associated with the Azores Current (AC). In most places the variability in the GEOSAT maps exceeds that of the CME model, with two notable exceptions: The spurious maxima in the model surface variability in the Irminger Sea and also near the Gulf Stream separation are not in agreement with the satellite observations.

It is evident from Figures 7.2c,d that both assimilation experiments successfully reproduce the spatial distribution of the rms surface variability in the GEOSAT data. Not only the NAC, but also the AC are indicated by regions of intensified variability, penetrating into the eastern basin. It is also consistent with the GEOSAT observations, that the position of the highest variability has moved from the Gulf Stream

separation region to the Gulf Stream extension. Moreover, the assimilation apparently can suppress the spurious maximum in the Irminger Sea. In general, the changes due to the assimilation are quite encouraging as far as the spatial distribution of surface height variability is concerned, whereas both assimilation experiments systematically underestimate the intensity of the rms surface height fluctuations. This is most pronounced for the re-initialization method (Fig. 7.2*d*), and will contribute to the above mentioned larger rms difference between the GEOSAT maps and the results of this assimilation scheme (Fig. 7.1). However, the rms variability does not contain much information about the representation of different physical processes. Neither does it discriminate between different wavenumber or frequency regimes, nor does it contain any phase information. We shall therefore proceed by examining the effect the assimilation has on actual model states, before investigating the contribution of individual wavenumber bands to the total rms surface variability.

Instantaneous SSH fields of both assimilation experiments are shown in Figures 7.3*c,d* for March 19, 1987, for the same area in the Gulf Stream extension that already served as example in section 5.3. These may be compared with the corresponding field of the model when no data are assimilated (Fig. 7.3*a*), and the combination of the model climatological mean surface height plus the SSH anomaly estimated from the 17-day period March 11-27 of GEOSAT data (Fig. 7.3*b*). In order to obtain a spatially complete GEOSAT estimate, this latter map again treats all data within the 17-day period as synoptic. (The actual map provided to the assimilation procedure just 32 timesteps (17 hours) ago uses an *e*-folding correlation decay time of 10 days and was shown together with the corresponding error map in Figure 6.5.)

While both assimilation methods can reproduce the position of the Gulf Stream meanders and also that of the cold core ring at 38°N, 56°W (surveyed by KRAUSS *et al.*, 1990) quite well, Figure 7.3 clearly reveals that the model fields are considerably smoother than the corresponding GEOSAT map. From the inspection of a number of regions at different times it seems to be a rather general phenomenon that the model is not able to follow the observations at small scales. This is further confirmed by the wavenumber spectra shown in Figure 7.4. For a number of 10° by 10° boxes annual averages of meridional spectra of the surface height anomalies are computed, treating spectra separated by 2° zonally and 26 days temporally as independent. To retain consistency with the model output, here the GEOSAT spectra are not computed along-track, but instead meridional spectra are derived from maps like the one shown in Figure 7.3*b*, also separated by 26 days and 2° zonally. Hence, the shortest wavelength that can be supported by the model grid (74 km meridionally) determines the

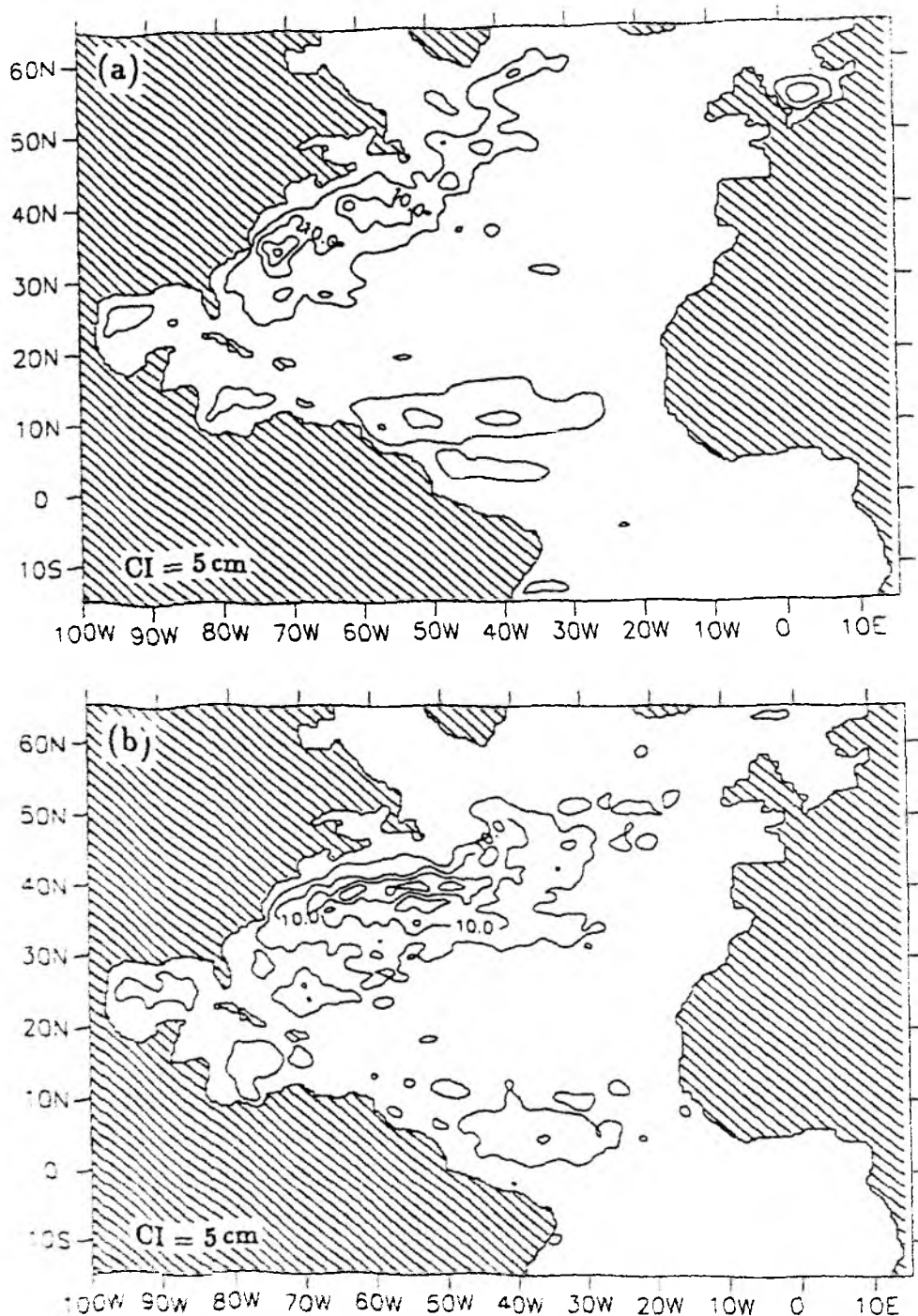


Figure 7.2: Variability of sea surface height computed from 14 maps separated by intervals of 26 days, for (a) CME model reference run with no data assimilated, (b) GEOSAT maps that treat all data within a 17-day period as synoptic.

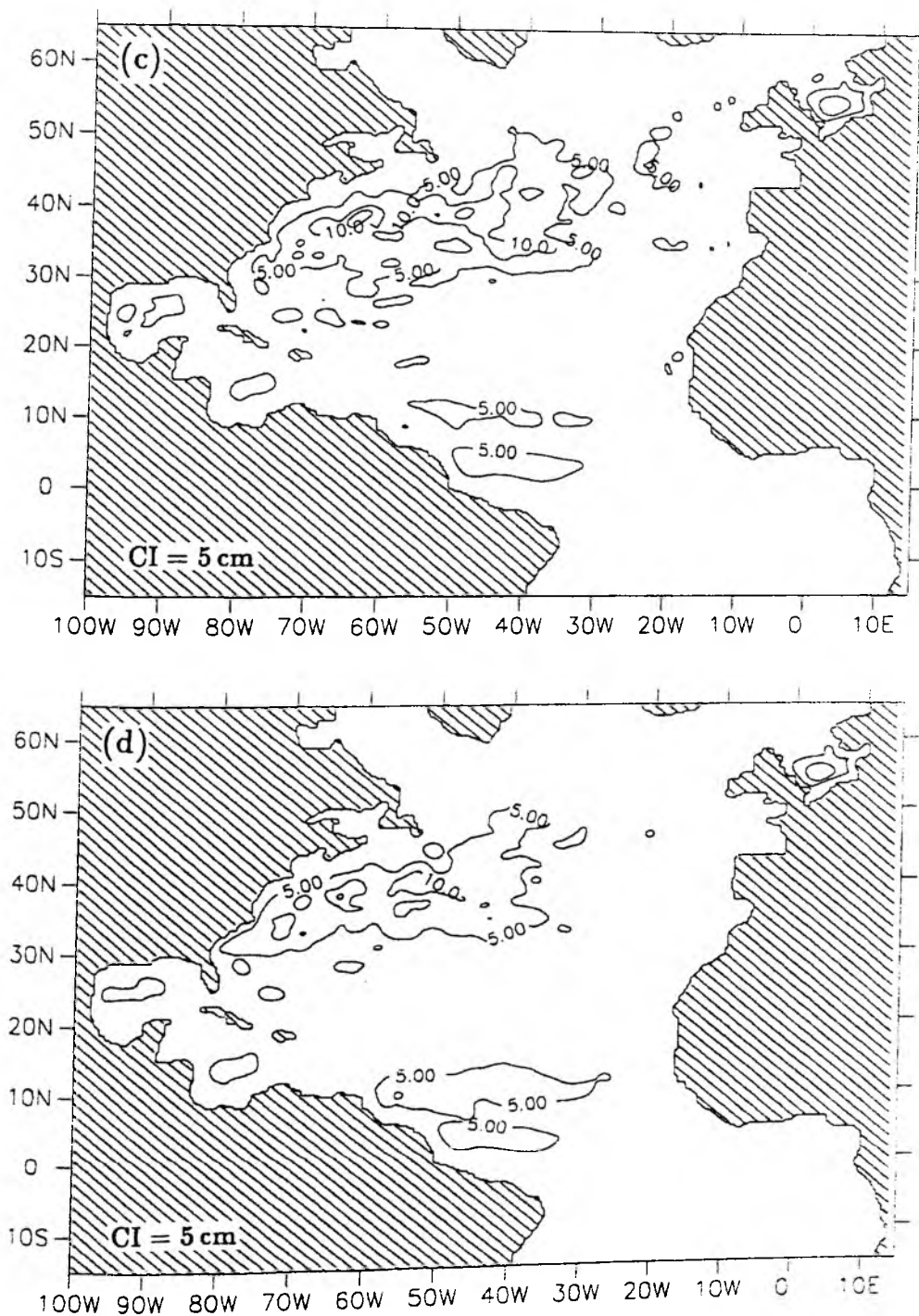


Figure 7.2: (continued): Sea surface height variability of the assimilation experiment using the extended nudging approach (c), and the re-initialization method (d).

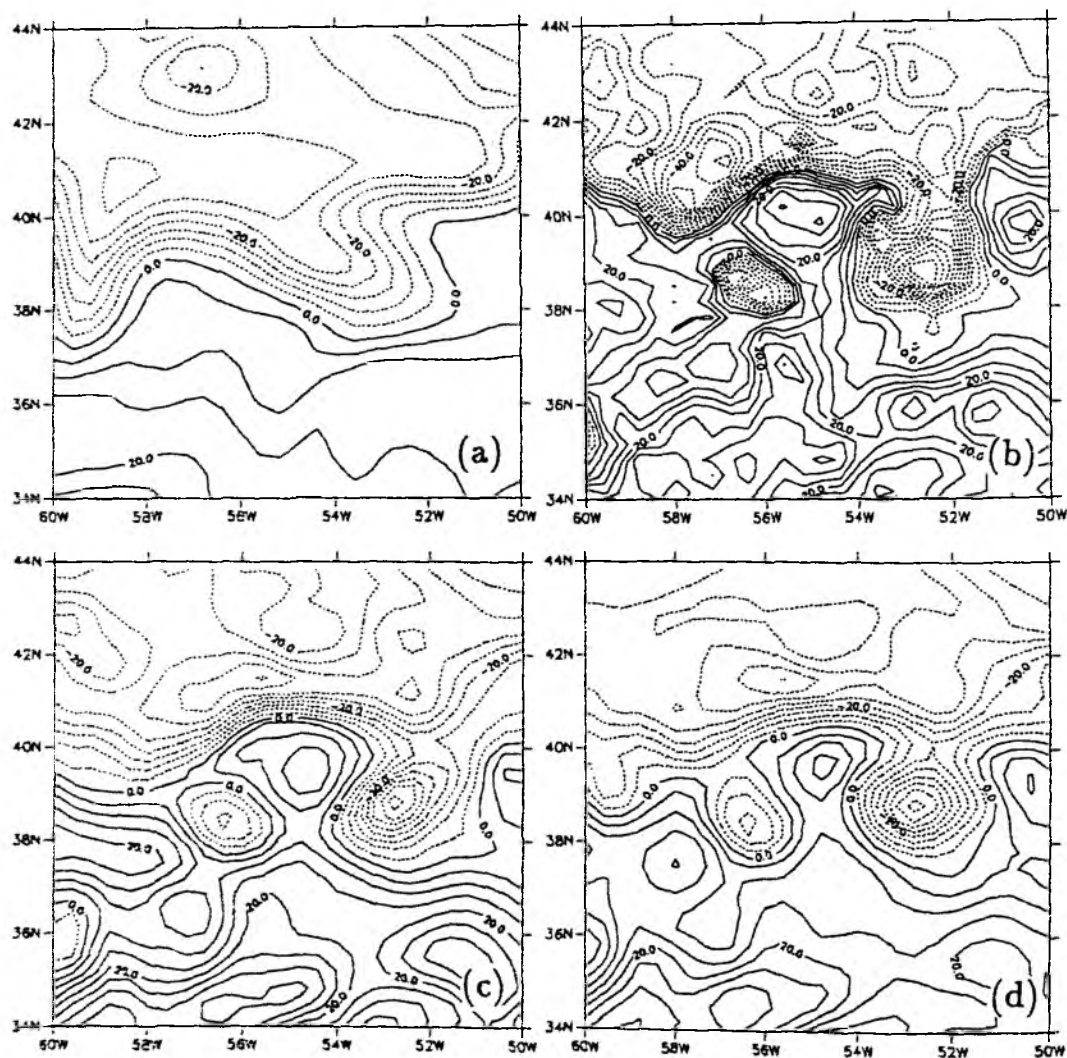


Figure 7.3: Sea surface height on March 19, 1987. (a) CME model when no data are assimilated, (b) model climatological mean plus SSH anomaly estimated from GEOSAT data of the 17-day period March 11-27. (c) Hindcast of the extended nudging experiment, and (d) of the re-initialization run. Contour interval is 5 cm.

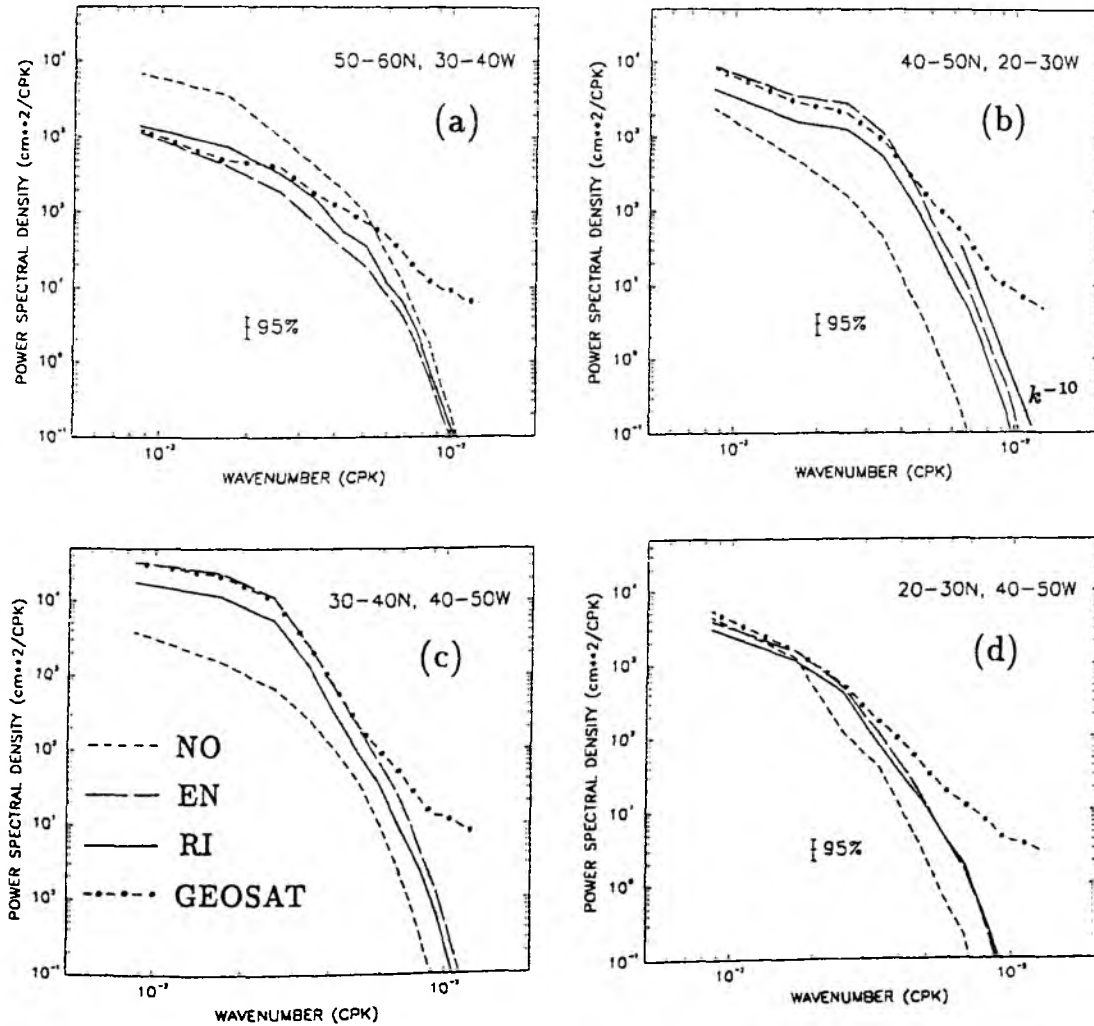


Figure 7.4: Meridional wavenumber spectra of sea surface height anomalies in 10° by 10° boxes. (a) Irminger Sea ($50^\circ - 60^\circ\text{N}$, $30^\circ - 40^\circ\text{W}$), (b) Northeastern Atlantic ($40^\circ - 50^\circ\text{N}$, $20^\circ - 30^\circ\text{W}$), (c) northern boundary of the subtropical gyre ($30^\circ - 40^\circ\text{N}$, $40^\circ - 50^\circ\text{W}$), (d) southern part of the subtropical gyre ($20^\circ - 30^\circ\text{N}$, $40^\circ - 50^\circ\text{W}$). Spectra shown refer to GEOSAT data, the model with no data assimilated (NO), the extended nudging (EN), and to the re-initialization experiment (RI).

cut-off wavenumber of these GEOSAT spectra. From Figure 7.4 it is obvious that the model spectra drop rapidly for wavelengths shorter than 150 – 200 km. As pointed out by BÖNING and BUDICH (1992) this is about the scale where, for a $1/3^\circ$ resolution model with biharmonic parameterization of lateral turbulent mixing, frictional control becomes dominant. Assuming geostrophically related surface fields, the dominance of biharmonic friction would result in a spectral relation close to k^{-10} . This can indeed be observed, being further consistent with the observation that for both assimilation experiments the basin averaged dissipation by horizontal friction turns out to be approximately twice as large as that of the reference run with no data assimilated – although the averaged kinetic energy remains roughly the same (see section 7.2). In this respect it is very likely that a model with higher horizontal resolution, that would allow shifting the frictional regime to smaller scales, can extract much more information from altimetric data – particularly on the mesoscale! For the present model resolution it is obvious that by smoothing the surface fields, the large viscosity of the model will contribute to the overall rms differences shown in Figure 7.1. However, the action of biharmonic friction does not explain the observed differences between the two assimilation experiments.

Further concentrating on wavelengths larger than 150 km we note that for all regions shown in Figure 7.4, both assimilation methods bring the model spectra closer to the GEOSAT data. At first sight the extended nudging approach seems to give somewhat better results than the re-initialization method, that tends to produce spectral energy densities in between those derived from the GEOSAT maps and the model run with no data assimilated. Nevertheless, this latter behaviour is exactly what one would generally expect for a scheme based on statistical interpolation between model forecast and observations. In contrast, the extended nudging experiment quite systematically *overestimates* the surface variability in places where the original model *underestimates* it (Fig. 7.4 b,c), and vice versa (Fig. 7.4 a). How can the results of this assimilation experiment exceed the variability level of the inserted data?

Actually, the discrepancy between the spectral energy densities referring to GEOSAT data and the CME model with no data assimilated, is a manifestation of incorrect model statistics. Any false estimation of these second order moments will have important consequences for the results of the extended nudging scheme, that uses the model statistics to infer estimates of the state variables from the observed SSH anomalies. Particularly the linear regression between sea surface height and temperature/salinity depends crucially on accurate statistics. As was pointed out in chapter 5, this is reason for some concern. We have already seen in the static snapshot example of section 5.3,

that applying the linear regression scheme to such surface observations that exceed the typical range of the original model values, may produce unrealistic water masses (Figure 5.2): Heavy water becomes too heavy, light water too light. A consequence is an overestimation of the vertical shear in the diagnosed geostrophic velocity field (Fig. 5.2 *m*), not consistent with the directly assimilated velocity field (Fig. 5.2 *l*). In a dynamical assimilation experiment, the effect of this inconsistency is felt in the period between successive assimilation cycles: Because of the large inertia of the deep layers, the surface currents will eventually spin up and correspondingly increase the variance of the surface height field. While biharmonic friction efficiently removes any small-scale perturbations, at scales larger than approximately 150 km it is only the following assimilation cycle that will reset the surface currents. For this reason, a dynamical application of the extended nudging approach of chapter 4 will tend to *overestimate* the surface variability in such regions where the original model *underestimates* it. The reverse argument also applies.

When comparing the two assimilation experiments with respect to the surface height fields, we conclude that although the results of the extended nudging approach show a somewhat better agreement with the GEOSAT observations, the slight overshooting of the spectral energy densities already reveals physical inconsistencies of this assimilation scheme due to the use of inaccurate statistics. These deficiencies will become more apparent when looking at water mass properties. However, before discussing this subject in greater detail, we shall for the moment turn towards the distribution of eddy kinetic energy, that is expected to closely follow that of the surface height variability.

7.2 Eddy Kinetic Energy

A major reason for assimilating altimeter data into the CME model is the expected increase of eddy energy to a more realistic level, needed to properly assess the role of mesoscale variability in the general circulation. For the present study, all deviations from the one-year mean of the respective model run are included in the term “eddy variability”. In particular, no distinction is made between mesoscale, seasonal or inter-annual fluctuations, as this would require a far longer integration time of the numerical experiments. A first impression of the distribution of eddy variability could already be obtained from the maps of rms surface variability (Figure 7.2). The corresponding eddy kinetic energy (EKE) close to the surface (at 50 m) is displayed in Figure 7.5. The results of the assimilation experiments clearly exhibit eastward reaching bands of high EKE values associated with the NAC and the AC, while spurious maxima close to the separation of the Gulf Stream and also in the Irminger Sea can efficiently be suppressed. Although the resulting EKE distribution pattern agrees quite well with that estimated from drifting buoys (RICHARDSON, 1983; KRAUSS and KÄSE, 1984) and also with along-track analysis of GEOSAT data (LETRAON *et al.*, 1990), modelled EKE levels are much too low: Observed maximum values in the Gulf Stream are $1800 \text{ cm}^2 \text{ s}^{-2}$ according to RICHARDSON, and values exceeding $2600 \text{ cm}^2 \text{ s}^{-2}$ have been deduced from the GEOSAT data. In the same region the model with no data assimilated barely reaches $250 \text{ cm}^2 \text{ s}^{-2}$, and the results of both assimilation experiments show maximum values of $400 \text{ cm}^2 \text{ s}^{-2}$. What is the reason of this still considerable underestimation of eddy kinetic energy by the model despite the assimilation of GEOSAT data?

As indicated by the SSH wavenumber spectra, shown in Figure 7.4, biharmonic friction becomes important only on scales smaller than 150–200 km. However, in the Gulf Stream region EKE spectra derived from GEOSAT data are “red” up to wavelengths of 300 to 400 km (LETRAON *et al.*, 1990). Hence, friction alone can not explain for the large energy gap between the GEOSAT EKE estimates and the results of the assimilation experiments. Instead, we shall argue that only a small amount of the estimated EKE distribution can actually be exploited by the assimilation procedure. This is essentially due to the large satellite ground track spacing of about 110 km at 40°N , which is further augmented by frequent data loss, mainly along the descending tracks crossing the Gulf Stream (see Fig. 6.1). This spacing is just about the lag of the first autocorrelation zero crossing (Fig. 6.2), and hence does not provide an adequate two-dimensional sampling of eddy activity in this region. Consequently, EKE levels re-computed from mapped SSH estimates (implying the assumption of isotropic SSH anomalies) can be

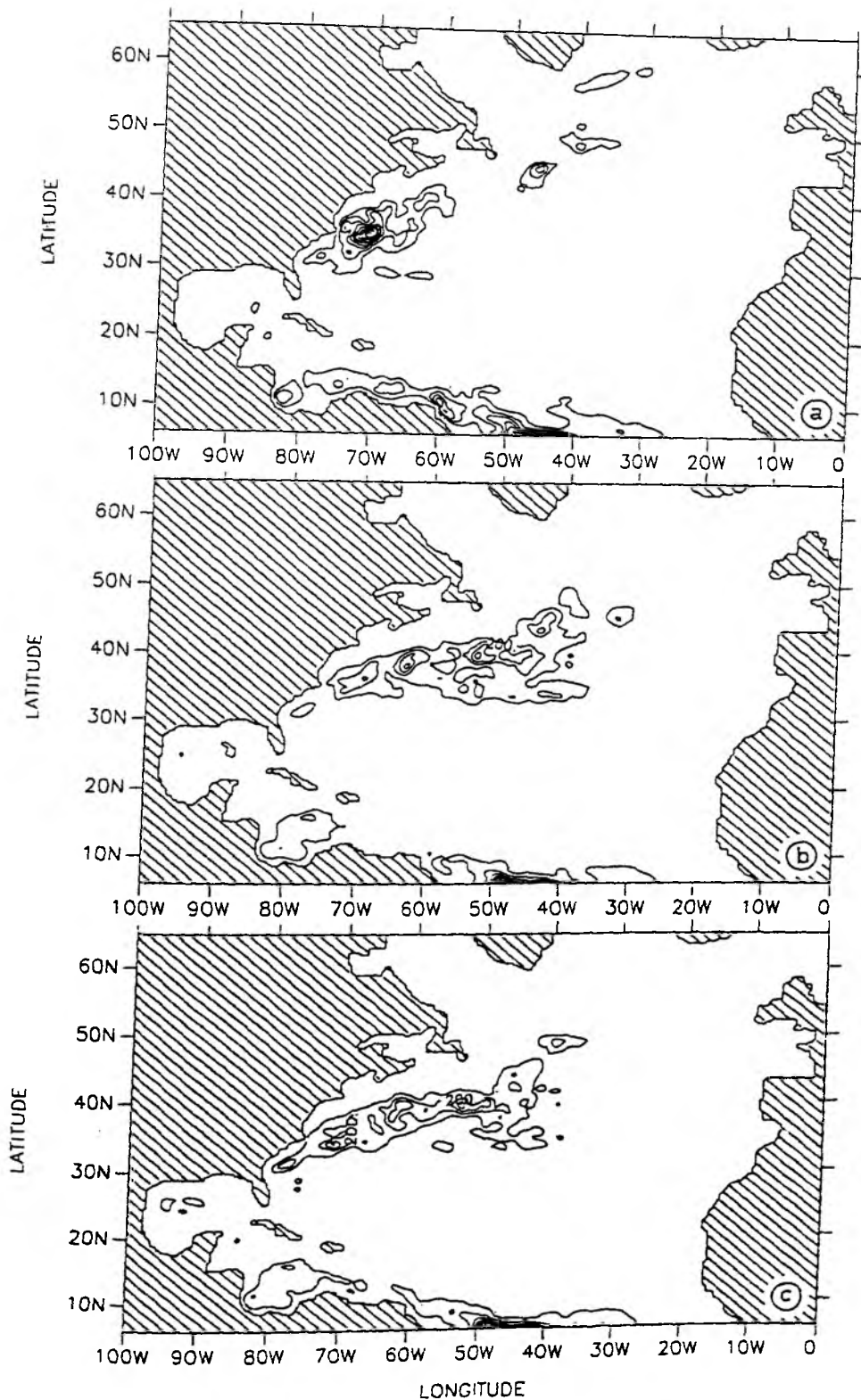


Figure 7.5: Eddy kinetic energy at a depth of 50 m. (a) no data assimilated, (b) extended nudging, and (c) re-initialization experiment. Contour interval is $100 \text{ cm}^2 \text{ s}^{-2}$.

significantly lower than EKE values obtained directly from the tracked data, assuming isotropy of velocity anomalies.

This is further illustrated by Figure 7.6 *a*, which shows the EKE distribution that would result from the one-year set of GEOSAT SSH maps, if all deduced surface velocities were accurate. Maximum values are $1600 \text{ cm}^2 \text{ s}^{-2}$, what already is considerably less than the $2600 \text{ cm}^2 \text{ s}^{-2}$ obtained by LETRAON *et al.* (1990) from two years of tracked GEOSAT data. Figure 7.6 *b* displays the energy corresponding to the weighted fraction $(1 - \tau)^2$ of the SSH gradients, where τ is the estimated relative error of the objectively analyzed SSH values. It is only this part of the eddy kinetic energy in the GEOSAT data, that the assimilation methods used in this study can immediately exploit. Any EKE values exceeding this level will have to be produced by the model itself, possibly as a dynamical reaction to the assimilated data. However, a comparison with Figure 7.5 reveals that the assimilation experiments in general do not develop significantly higher eddy activity than directly provided by the inserted data.

Any attempt to further increase the eddy kinetic energy of the CME model by assimilating altimeter data, using one of the techniques proposed in this study, will therefore to a large degree depend on more accurate surface velocity (and hence EKE) estimates. In this respect, the dense ground track pattern resulting from the 35-day repeat orbit of the European satellite ERS-1 may provide some advantages over the 10-day repeat cycle of the joint altimetric TOPEX/POSEIDON mission of the United States and France.

Having so far concentrated on eddy variability at the ocean surface, we shall now examine to what extent the assimilated surface information can penetrate into the ocean interior. Exemplarily, a zonal section of EKE at 48°N is presented in Figure 7.7. As described in greater detail by TREGUIER (1992), who compared the same CME model section with the TOURBILLON and TOPOGULF experiment moorings along 48°N , the eastward decrease of EKE is too abrupt in the model. For reference, these mooring results, obtained by COLIN DE VERDIÈRE *et al.* (1989), have been included in Figure 7.7 *d*. Between 30°W and 15°W EKE values of 40 to $60 \text{ cm}^2 \text{ s}^{-2}$ were measured at a depth of 600 m , with values larger than $100 \text{ cm}^2 \text{ s}^{-2}$ at 300 m , whereas the model hardly exceeds $2 \text{ cm}^2 \text{ s}^{-2}$. This may be compared with the results of the two assimilation experiments, shown in Figures 7.7 *b, c*. In both cases, EKE has considerably increased east of 30°W , with values generally exceeding $10 \text{ cm}^2 \text{ s}^{-2}$ in the upper 250 m . Interestingly, the two assimilation experiments reveal rather different vertical gradients in the EKE distribution. While the extended nudging scheme tends to produce higher surface values than the re-initialization method, the reverse is true at depths

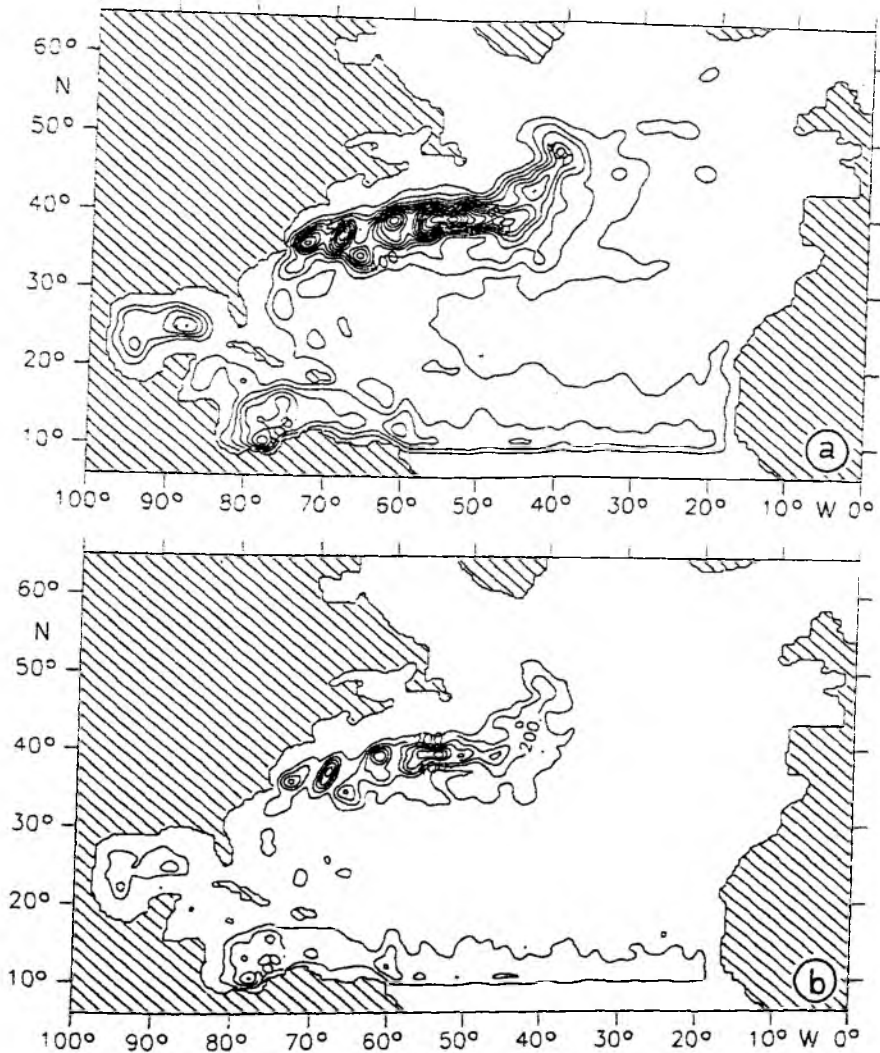


Figure 7.6: (a) Eddy kinetic energy computed from the one-year five-daily set of GEOSAT SSH maps. The inhomogeneous pattern in the Gulf Stream extension is due to frequent data gaps along the descending tracks (see also Fig. 6.1). (b) EKE that would correspond to the weighted fraction $(1 - r)^2$ of the estimated surface velocities. Contour interval is $100 \text{ cm}^2 \text{ s}^{-2}$.

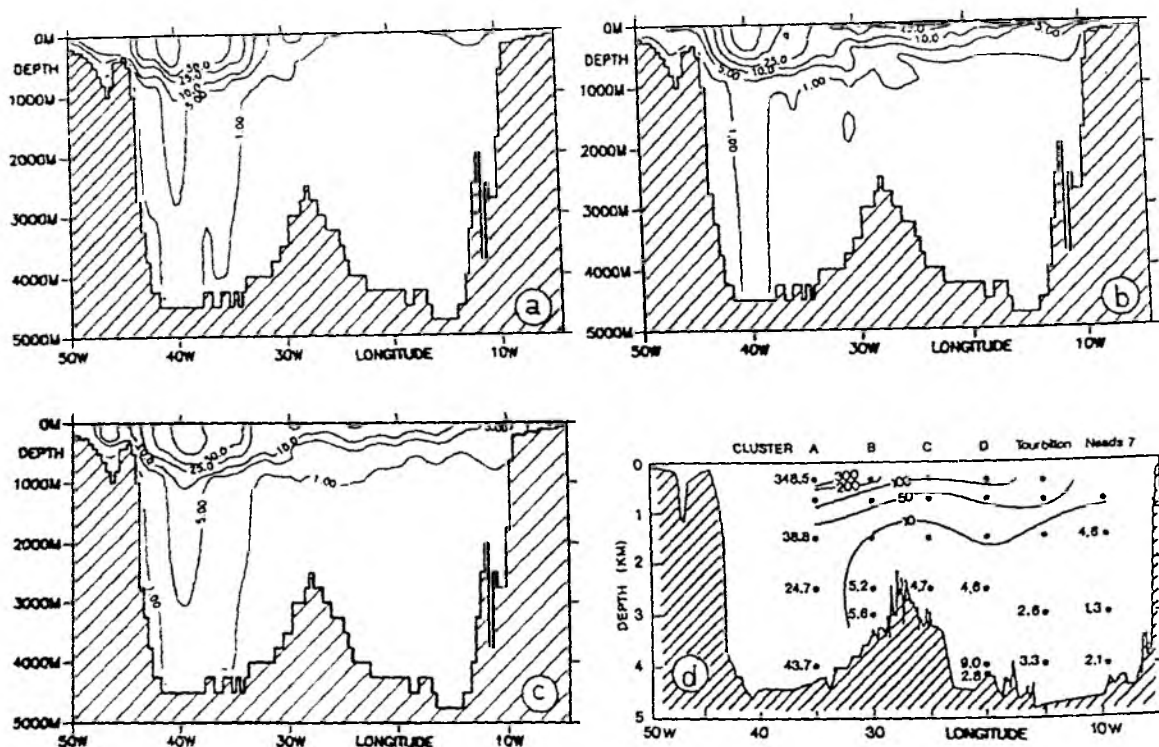


Figure 7.7: Zonal section of *EKE* at 48N, (a) no data assimilated, (b) extended nudging, and (c) re-initialization experiment, (d) observed values reproduced from COLIN DE VERDIÈRE et al (1989). Isolines in (a), (b), and (c) are at 1, 5, 10, 25, 50, 100 $\text{cm}^2 \text{s}^{-2}$.

below 500 m, where the extended nudging scheme produces EKE values even lower than those of the original model with no data assimilated!

To understand this pronounced difference between the two assimilation experiments, we again have to bear in mind that a systematic underestimation of eddy variability by the original model automatically implies inaccurate model statistics. A typical result of combining these statistics with the vertical projection schemes of the two alternative assimilation methods has been given by the static example of section 5.3. Any overestimation of the vertical gradients of density anomalies will essentially produce too intense geostrophic near-surface currents. For this reason the vertical projection scheme of the extended nudging approach can lead to higher surface eddy activity than the dynamically balanced assimilation of the re-initialization method. However, it does not explain the reduced EKE values at greater depths, that actually result in the basin averaged kinetic energy of the extended nudging experiment (annual mean

of $13.1 \text{ cm}^2 \text{ s}^{-2}$) being lower than that of the reference experiment with no data assimilated ($13.6 \text{ cm}^2 \text{ s}^{-2}$). (The average kinetic energy of the re-initialization experiment amounts to $14.2 \text{ cm}^2 \text{ s}^{-2}$.)

The fact that despite assimilating GEOSAT data – with a higher energy level (Fig. 7.6) than the original model (Fig. 7.5a) – the total kinetic energy of the extended nudging experiment decreases, deserves some attention. Actually, the vertical extrapolation scheme used by the extended nudging method to produce the required “pseudo” observation is essentially aimed at minimizing the *error* variance of the state estimate. In the absence of perfect correlations with the observed surface variables, this implies a reduction of the estimated *process* variance. Therefore, even very cautious nudging of the model state towards such “pseudo” observations will eventually result in some energy deficit. In contrast, the re-initialization method uses just the estimated model-data misfit to update the actual model state. Correspondingly, the variance of the update increments may be underestimated, but that of the model – and hence its energy – is not necessarily reduced.

A final example illustrating the effect of the assimilation on the model’s representation of eddy variability is given by a meridional section along 30°W . The EKE distribution of the CME model with no data assimilated is displayed in Figure 7.8a. It reveals maximum values at 60°N , associated with the unrealistic band of high eddy activity in the Irminger Sea. In contrast, the eddy signal of the NAC at 51°N and also of the AC at 34°N is rather weak, with maximum values of only $30 \text{ cm}^2 \text{ s}^{-2}$. This is an order of magnitude smaller than surface drifter EKE levels as for example reported by KRAUSS and KÄSE (1984)! At 30°W they obtained more than $400 \text{ cm}^2 \text{ s}^{-2}$ for the NAC, and about $150 \text{ cm}^2 \text{ s}^{-2}$ for the AC.

At this 30°W section, both assimilation methods can approximately double the surface EKE associated with the AC. Again, the re-initialization method yields a somewhat deeper penetration of the enhanced surface variability. On the other hand, the spurious maximum at 60°N , not confirmed by observational evidence, has completely vanished only in the extended nudging experiment. An interesting feature can be deduced from the EKE distribution between 45°N and 54°N . While the model with no data assimilated produces a single EKE maximum in the NAC region at 51°N , both assimilation experiments suggest the existence of at least two bands of intensified eddy variability, with values exceeding $50 \text{ cm}^2 \text{ s}^{-2}$. A southern branch of enhanced variability between 45°N and 48°N is separated from the northern maximum at 51°N by a quiet region with values less than $20 \text{ cm}^2 \text{ s}^{-2}$. This pattern of multiple local

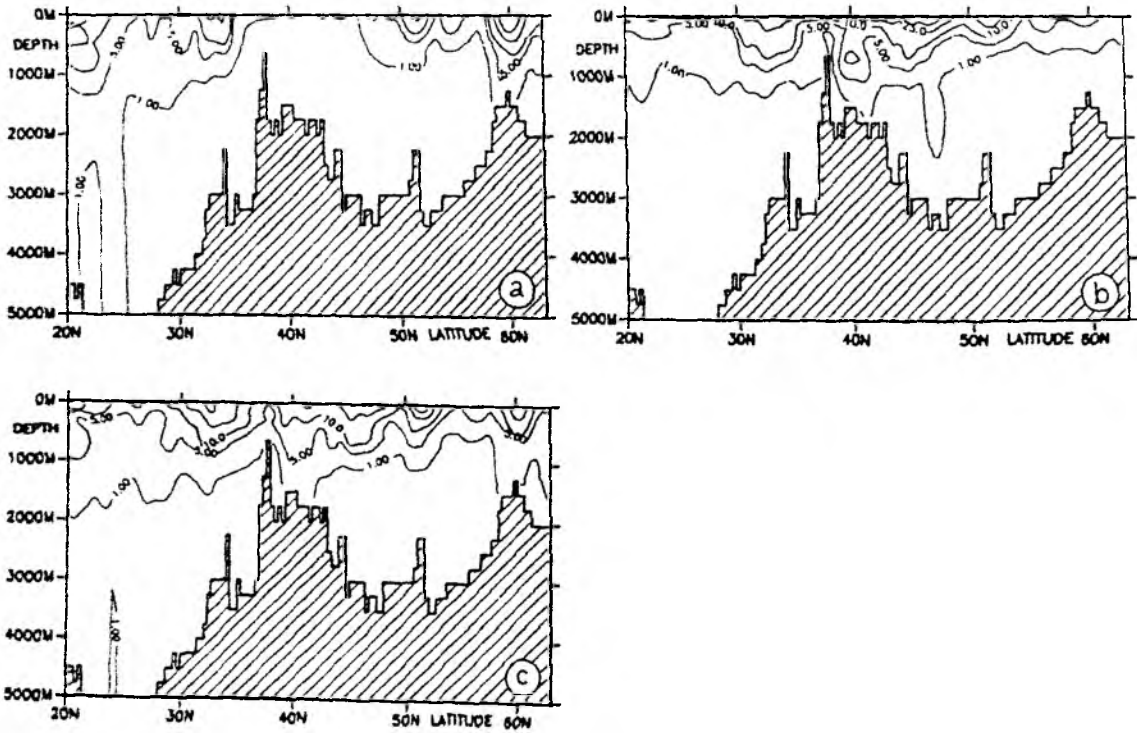


Figure 7.8: Meridional section of EKE at 30°W, (a) no data assimilated, (b) extended nudging, and (c) re-initialization experiment. Isolines are at 1, 5, 10, 25, 50, 100 cm^2s^{-2} .

EKE maxima agrees very well with observational findings reported by SY *et al.* (1992). Analyzing a number of hydrographic sections, they concluded that the NAC crosses the MAR in form of individual current branches. Only the northernmost branch, associated with the subarctic front, was found to be topographically fixed at the Gibbs Fracture Zone at 52°N, while the remaining current bands varied in number, location and intensity. Interannual fluctuations of these current bands, and in particular a two-banded structure in the year 1987, were also reported by HEYWOOD *et al.* (1994), who inferred surface EKE distributions for different years from altimetric data sets.

In conclusion we therefore note that both assimilation methods quite successfully change the spatial EKE distribution with an accuracy that allows monitoring of inter-annual variability in the location of frontal systems. However, the resulting vertical EKE profiles are rather distinct. The re-initialization method is not only more robust to errors in the model statistics, but it also avoids any systematic underestimation of

the process variance. While in the extended nudging experiment the former leads to an overestimation of the surface variability, the latter produces too low energy levels at depth. In this respect, our comparison of the two alternative assimilation schemes has already revealed some advantages of the re-initialization method, which, after all, will become more pronounced in the following sections.

7.3 Hydrography

We have just seen that the two assimilation experiments produced rather distinct vertical EKE distributions, indicating a different baroclinic response of the ocean model to the inserted data. This was at least partly ascribed to the different treatment of temperature and salinity by the two alternative assimilation schemes. It is the subject of the present section to examine the resulting changes in the hydrography in greater detail, and to further clarify their different dynamical implications.

Actually, statistical analysis of large records of expendable bathythermograph (XBT) data considerably contributed to the emerging picture of the global distribution of eddy variability long before the advance of sophisticated satellite technology (DANTZLER, 1977; EMERY, 1983). Fluctuations in the baroclinic current field are related to vertical displacement of isopycnal surfaces, corresponding to temperature variations on isopleths. In this respect, any changes in the temperature rms variability at a fixed depth level will indicate baroclinic flow changes due to the assimilation. Figure 7.9 compares the rms temperature variability at a depth of 460 m, obtained from XBT data by EMERY (1983), with the variability found in the one-year CME model experiments. The XBT data reveal maximum standard deviations of 3°C in the Gulf Stream and also east of Newfoundland. These maxima are quite well reproduced by the model without data assimilation (Fig. 7.9*b*), although – similarly to the surface EKE – the eastward decrease of eddy variability is too rapid. Not consistent with the observations are the high rms values in the Irminger Sea. The results of both assimilation experiments show a more realistic eastward penetration of the zones of high variability, associated with the NAC and the AC. Nevertheless, the extended nudging scheme tends to underestimate the overall baroclinic eddy activity at this depth. Temperature standard deviations barely exceed 1°C. In contrast, the re-initialization scheme achieves rms values larger than 2.5°C, agreeing much better with the XBT data. This corresponds very well to our findings, that EKE intensities below a shallow surface layer are generally too low for the extended nudging experiment. It can again be attributed to

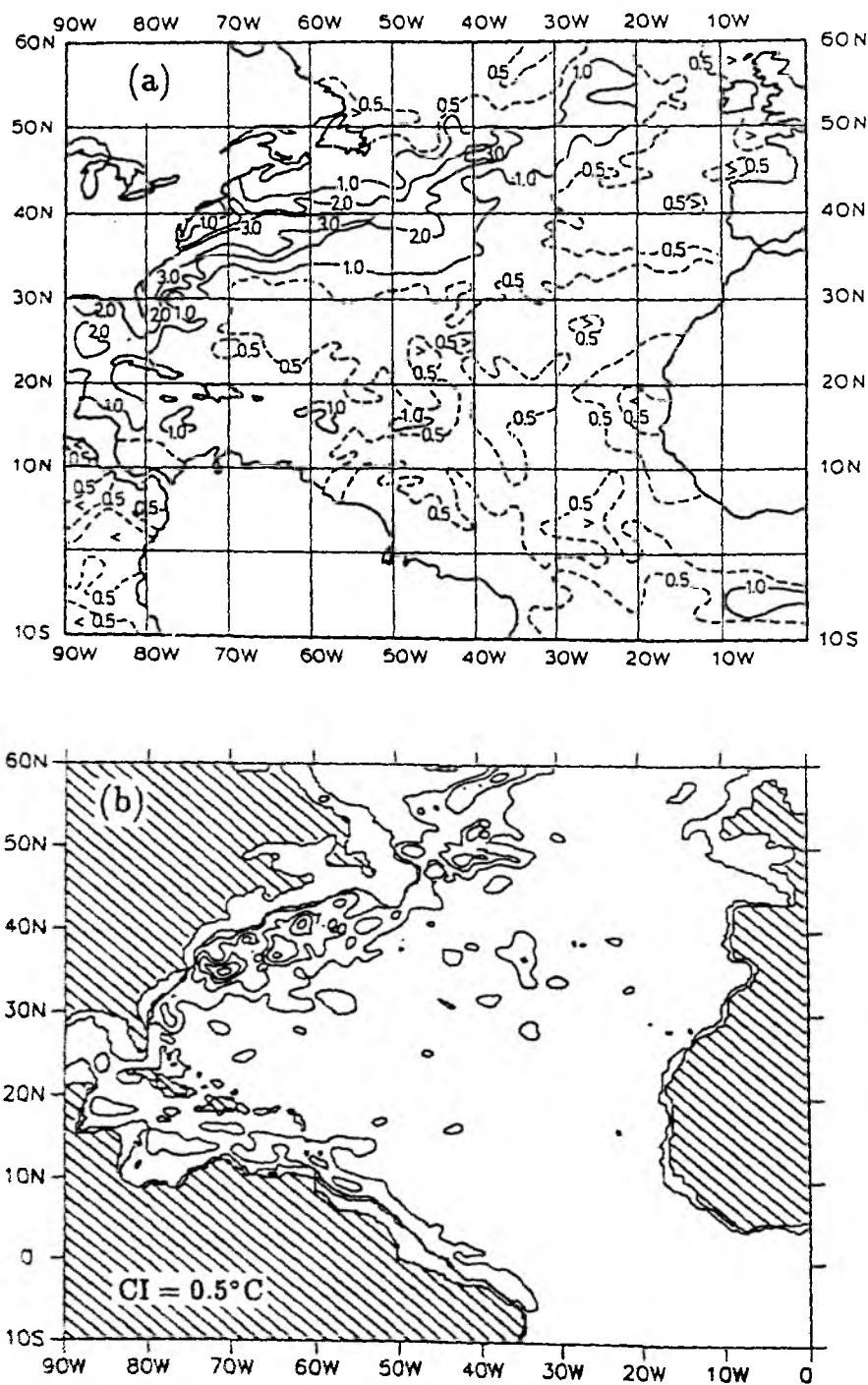


Figure 7.9: Standard deviation of temperature at 460m. (a) analysis of XBT data by EMERY (1983), (b) CME model with no data assimilated.

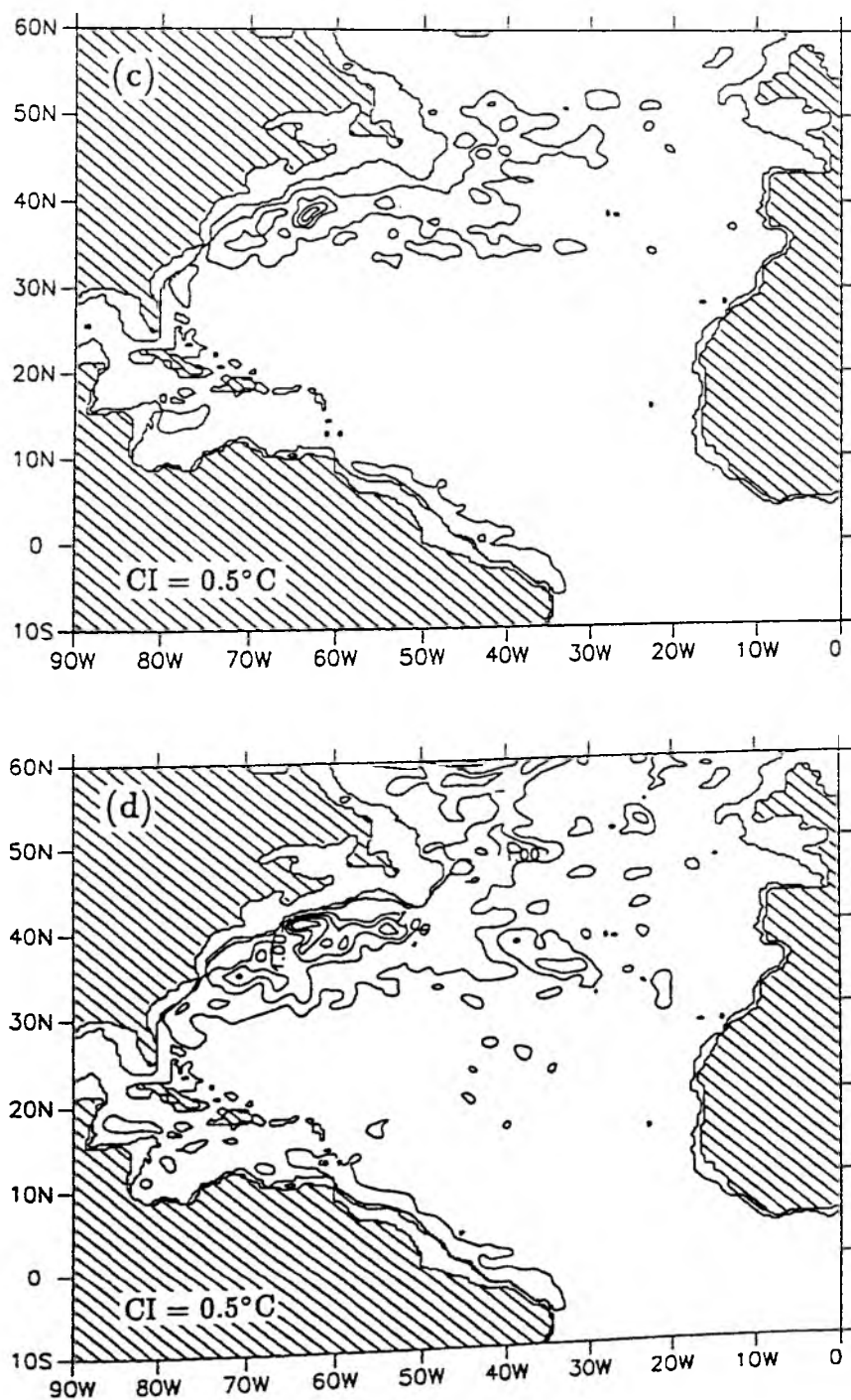


Figure 7.9: (continued): Standard deviation of temperature at 460 m. (c) result of extended nudging assimilation experiment, (d) re-initialization experiment.

the tendency of the vertical linear regression scheme to systematically underestimate the process variance of the estimated “pseudo” observations.

Because the spatial distribution of temperature rms variability does not only change, but in general improves when compared with the independent XBT data set, one may conclude that the treatment of baroclinic effects by either assimilation scheme is not completely unreasonable. However, the rms variability is just a statistical property and does not indicate whether the assimilation of altimeter data can in fact lead to an improved description of the actual hydrographic state of the ocean. To investigate this matter in greater detail, we shall exemplarily analyze an instantaneous state of the numerical ocean model. The state chosen corresponds to the static example of section 5.3, that was basically set up to test the alternative vertical projection schemes of the two assimilation methods. Concentrating on the effect of *dynamical* assimilation, we shall now examine the hindcast, produced by the two assimilation experiments for March 19, 1987. Analogously to Figure 5.2 of the static example, density, temperature and zonal velocity are plotted along a meridional section at 56°W in Figure 7.10 together with the diagnosed velocity that would balance the density fields geostrophically. The corresponding SSH maps were shown in Figure 7.3 of section 7.1.

Recall that according to the hydrographic survey taken by KRAUSS *et al.* (1990), this meridional section should contain a cold core ring near 38°N, just south of the Gulf Stream front at about 41°N. This cold core ring is indeed nicely reproduced by the re-initialization scheme (Fig. 7.10 *f*). Agreeing very well with the observations reported by KRAUSS *et al.* (see Fig. 5.3), the raised dome structure in temperature and density extends down to the bottom of the thermocline. In contrast, the cold core ring of the extended nudging hindcast is much less pronounced. Instead, unrealistically high temperatures ($> 23^{\circ}\text{C}$) are simulated at the offshore side of the Gulf Stream front, corresponding to densities much too low ($\sigma_0 < 24.5$). Clearly, the hydrographic structure of the re-initialization hindcast is the more realistic one.

To get an impression of the different results produced by dynamical and static assimilation, we can now directly compare Figure 7.10 with the static example displayed in Figure 5.2. Most of the differences in the mixed layer can at least partly be ascribed to the action of restoring surface boundary conditions. However, any differences below the mixed layer must be due to the capability of the dynamical assimilation procedure to accumulate and distribute information. For the chosen example section, this has the greatest effect on the thermocline structure of the re-initialization experiment, especially in the region of the cold core ring. Actually, only this observed difference between static and dynamical assimilation justifies the effort of combining altimeter

data with a complex numerical model. The role of the model is not merely restricted to interpolate successive state estimates in time, but it plays an active part in extracting information from the data that could not be gained from individual static analysis steps.

Pronounced differences between dynamical and static assimilation also occur in the velocity field. The fact that both dynamical assimilation experiments produce more intense currents than the corresponding static assimilations, can also be attributed to the accumulation of information over a number of assimilation cycles. In contrast to the static estimates (Fig. 5.2 *g, l*), the profiles of zonal velocity obtained from the dynamical assimilation experiments differ considerably from each other (Fig. 7.10 *g, l*). In accordance with the density fields, the current velocities of the re-initialization experiment reveal weaker vertical gradients, with the 30 cm s^{-1} contour reaching a depth of almost 500 m, whereas such intense current velocities are confined to the upper 150 m in the extended nudging hindcast. Interestingly, the differences between the simulated zonal velocity and the one diagnosed from the corresponding density field (Fig. 7.10 *h, m*) show about the same behaviour for both dynamical assimilation experiments. This is in sharp contrast to the strongly unbalanced fields emerging from the static assimilation using the statistical projection scheme of the extended nudging method (Fig 5.2 *m*). In the dynamical assimilation experiment, the finite width of the nudging window allows geostrophic adjustment processes to take place and to reduce the imbalances that would result from the static assimilation to a level acceptable for a proper model evolution. In this respect, the intention of the nudging approach to let the model dynamics find a state dynamically consistent with the inserted data, is realized. However, the unrealistic water mass properties produced by the statistical vertical extrapolation scheme remain a serious drawback of the extended nudging method.

A last example will be given to further underline that the re-initialization method is much more robust to imperfections in the model statistics. In Figure 7.11 the annual temperature cycle at 40°N , 30°W is displayed. Unlike most regions in the eastern North Atlantic, at this position the SSH rms variability of the original CME model (2.6 cm) is still of the same order of magnitude as the rms variability observed by GEOSAT (4.2 cm). Nevertheless, the extended nudging scheme produces a completely unrealistic annual cycle in the temperature field. The seasonal cycle in the mixed layer has almost disappeared, while the deep isotherms undergo large excursions.

This not really satisfactory phenomenon is caused by the fact that different frequency bands contribute to the rms variability of the model on the one hand, and of the GEOSAT data on the other hand. Some indication about the quality of the

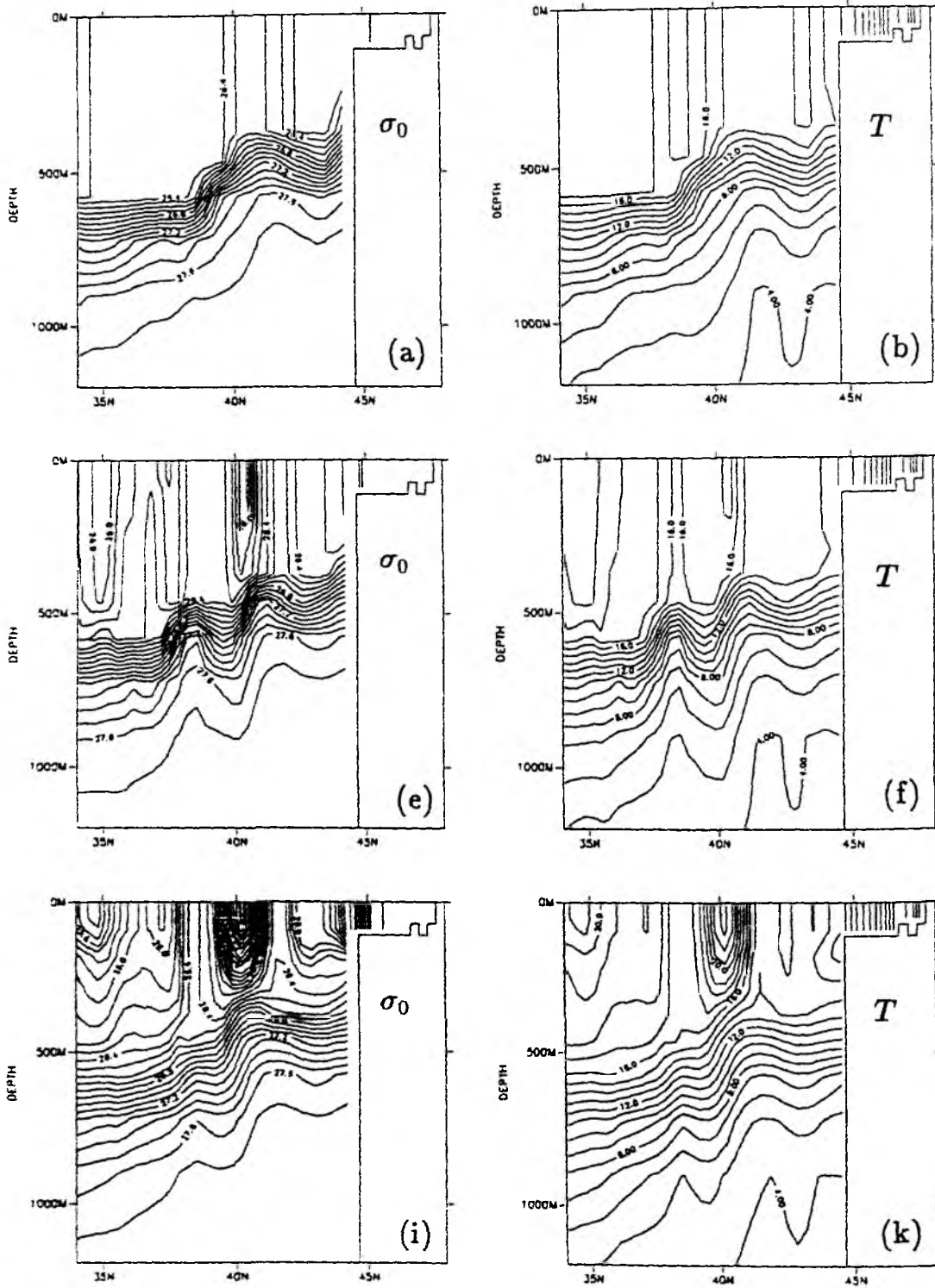


Figure 7.10: Meridional sections at 56°W . Top row: Forecast fields for March 19, 1987, by the CME model run with no data assimilated. Middle: Hindcast of the re-initialization experiment. Bottom: Hindcast of the extended nudging experiment. (a),(e),(i) show sections of potential density σ_0 , (b),(f),(k) potential temperature,

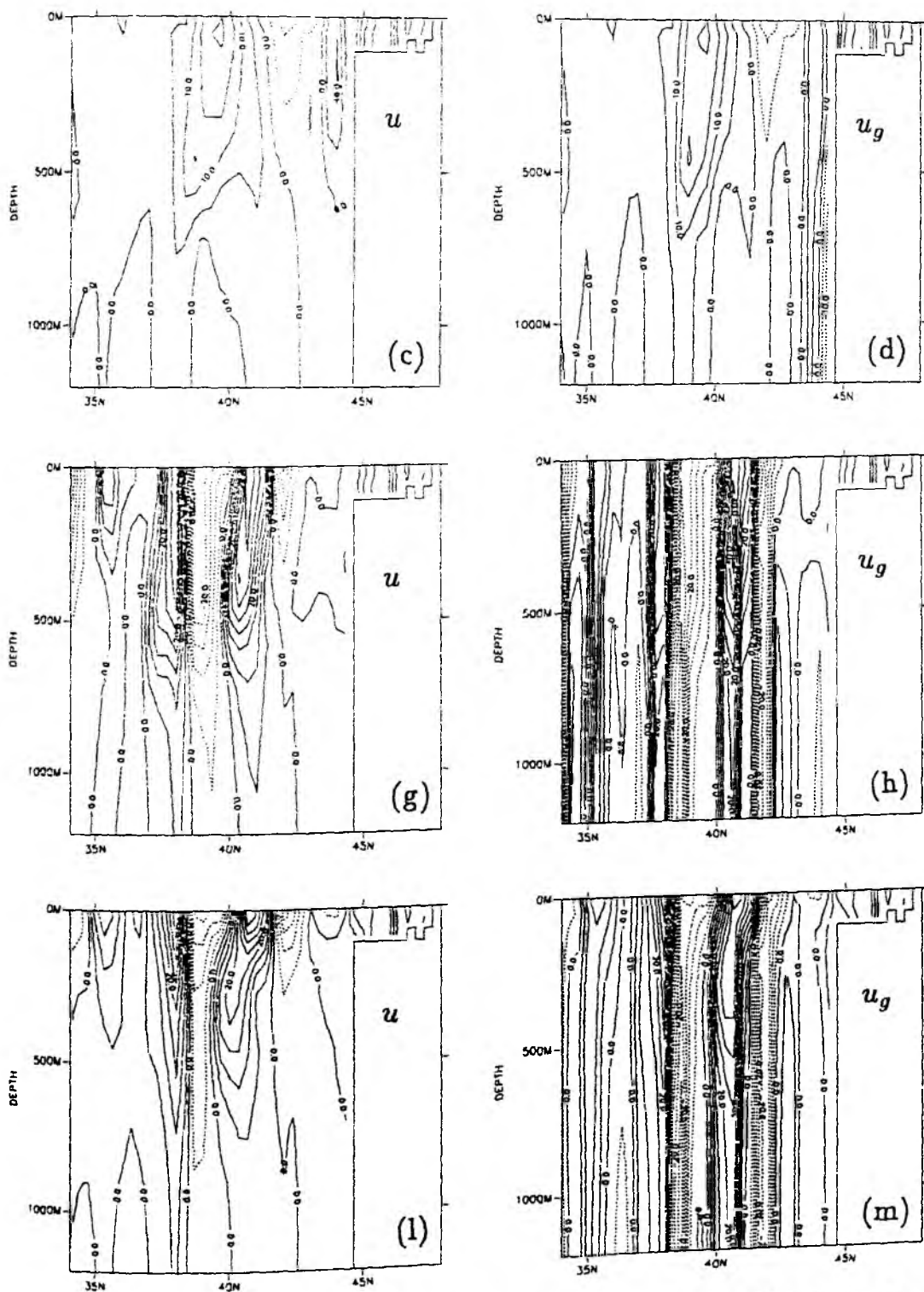


Figure 7.10: (continued): (c),(g),(l) zonal velocity, and (d),(h),(m) the diagnosed zonal velocity that would balance the density field of the first column geostrophically, with the surface velocity fixed by the values of the third column. Contour intervals are 0.1 for σ_0 , 1°C for temperature, and 5 cm s^{-1} for velocity.

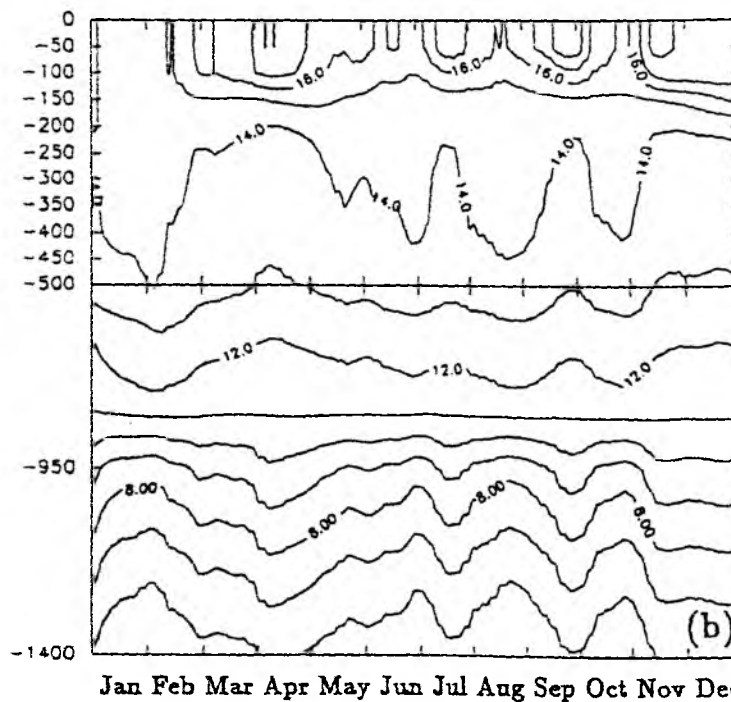
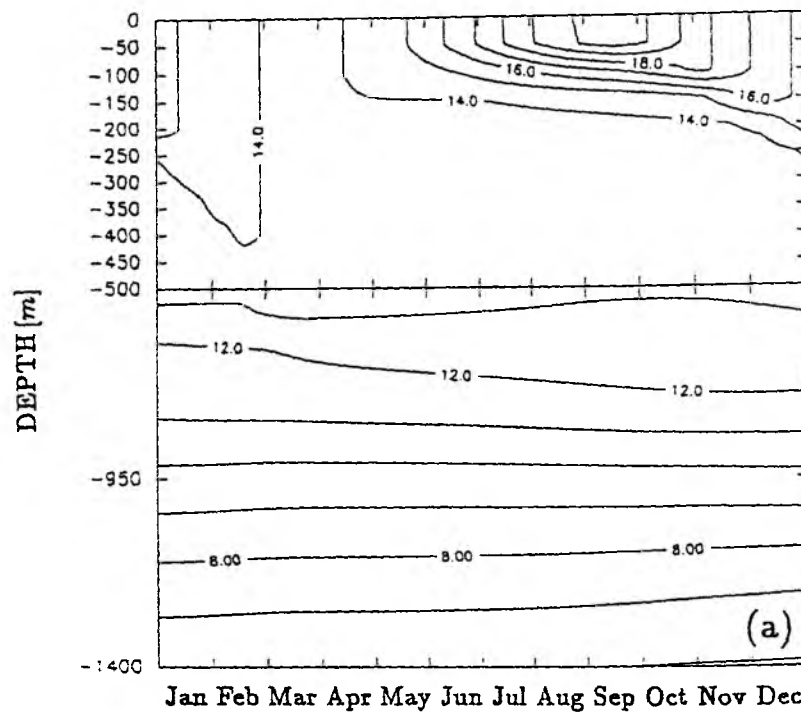


Figure 7.11: Annual cycle of temperature at 40°N, 30°W, (a) CME model with no data assimilated, (b) extended nudging experiment.

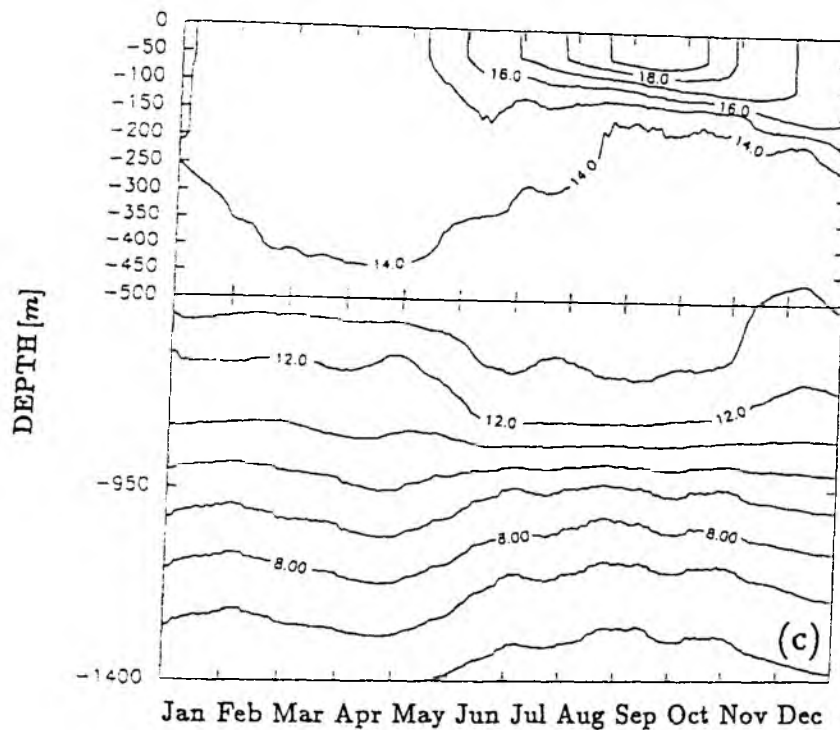


Figure 7.11: (continued): (c) *annual cycle of temperature at 40°N, 30°W, produced by the re-initialization experiment.*

statistical regression parameters obtained from three model years, can be found in the scatter diagrams, that for exactly this station were shown in Figure 4.7. The linear regression, relating temperature at 900 m to sea surface height, is dominated by the seasonal cycle superimposed on a warming trend of about $0.6^{\circ}\text{C}/\text{year}$, associated with a sea level rise of 3 cm/year, that is probably part of multi-year period oscillation. For the same region, the GEOSAT data show a much larger contribution of mesoscale fluctuations to the overall rms variability (STAMMER, 1992). Applying the model statistics to these data results basically in the model switching to summer (winter) conditions of the third (first) model year used to compute the statistics, whenever GEOSAT observes a positive (negative) SSH anomaly.

Shown in Figure 7.11 c is the much better behaved annual cycle that results from the re-initialization experiment. In contrast to the extended nudging scheme, the seasonal cycle of the mixed layer is well conserved. The largest temperature changes due to the assimilation of mesoscale SSH fluctuations occur below the mixed layer, where the vertical shear in the velocity field is largest. By using correlations between velocity components at different depth levels only, the re-initialization method obviously is

more robust to inaccurate model statistics. A further conceptual advantage of this assimilation scheme is the use of the T-S profile of the current model forecast to compute the new hydrographic structure at each assimilation step (section 5.2). In principle this allows for systematic deviations from the model climatology, as long as they help to keep the model evolution on track with the observations. While deviations from the climatological mean that lead to a wrong SSH forecast will be eliminated at the following update step, changes in the model hydrography that are dynamically consistent with the subsequent evolution of the observed SSH anomaly field, can persist and eventually change the climatological mean state. An example of this phenomenon, illustrated by the annual cycle of Figure 7.11 c, is the intensification of the vertical temperature gradient in the mean thermocline in June. Its structure can persist for about five months despite considerable eddy activity, e.g. seen in the results from the extended nudging experiment (Fig. 7.11 b), which uses the model climatology as reference at each assimilation step and thereby causes the state variables to essentially fluctuate around their climatological mean. Hence, applying the extended nudging scheme strongly inhibits any significant changes in the model climatology. The re-initialization method, on the other hand, does in principle allow the model to change its mean state, and it is the objective of the next section to identify such changes.

7.4 Changes in the Model Climatology

As far as the re-initialization method is used to assimilate satellite measured SSH anomalies, at least two principal mechanisms that may potentially lead to changes in the model *mean* can be imagined:

First, by correcting just for the misfit between model forecast and observations, the re-initialization scheme supports changes in the mean flow that improve the subsequent evolution of SSH anomalies (e.g., advection of eddies by the mean flow). However, in this study the GEOSAT measurements are referenced to the mean surface height provided by the model climatology. For this reason only such changes can occur that essentially conserve the climatological mean surface height. Nevertheless, by taking into account the estimated error variance of the mapped SSH anomalies, the re-initialization scheme does not require the assimilation model to fit the observations exactly. As will become evident later, the existence of this error margin allows for small changes even in the mean surface height.

Second, a significant amount of water mass formation is believed to take place in

the center of cyclonic anomalies or at rather unsteady frontal systems like the NAC. It was shown in the previous section (Fig. 7.10) that the re-initialization method could quite accurately reproduce position and structure of rings and fronts. In this respect, the temporal and spatial setting of intermittent water mass formation is likely to be controllable by the assimilation of altimeter data. With temperature and salinity of the newly formed water essentially being fixed by the atmospheric conditions in the formation area, the quality of the water masses produced by the assimilation experiment then depends crucially on accurate surface boundary conditions. Because the re-initialization scheme explicitly conserves temperature and salinity on isopycnals, subsequent assimilation steps will not alter these properties once the water mass has been subducted.

Figure 7.12 displays the changes in the annual mean temperature at a depth of 180 m, that result from the assimilation of SSH anomalies using the re-initialization method. Shown is the difference between the re-initialization experiment and the model run with no data assimilated. A large region of increased temperature is found slightly south of the NAC, extending from the Grand Banks far into the eastern basin. This is accompanied by a temperature decrease to the north of the Gulf Stream and also north of the NAC. In general, these changes tend to strengthen the cross-frontal density gradients and hence are expected to be associated with intensified mean currents in this region. This is indeed confirmed by the observed changes in the mean near-surface velocity field, displayed in Figure 7.13. For a broad region slightly southeast of the mean position of the NAC in the original CME model (Fig. 7.13b), there is a general increase in current velocities towards the northeast. Typically, changes in the mean surface currents amount to 10 to 20 cm s⁻¹. Note in particular the zonal band of increased northeastward current velocities at about 46°N, 30°W. Actually, the distribution of eddy kinetic energy both deduced from GEOSAT data (HEYWOOD *et al.*, 1994) and also emerging from the assimilation experiments (Fig. 7.8), already suggested that in the year 1987 a current band of the NAC crossed the Mid-Atlantic Ridge at about 46°N. While this conclusion so far was solely based on the observed increase in eddy variability, the manifestation of this current branch also in the annual mean flow provides further evidence to complete this picture.

An improvement of the mean currents is also believed to occur in the Irminger Sea. As an effect of the assimilation, the unrealistic and rather strong northeastward current produced by the original CME model southeast of Cape Farewell (Fig. 7.13b) is systematically weakened in the re-initialization experiment. Additional changes in the mean flow are seen in the Azores Current region, opposing the eastward flowing

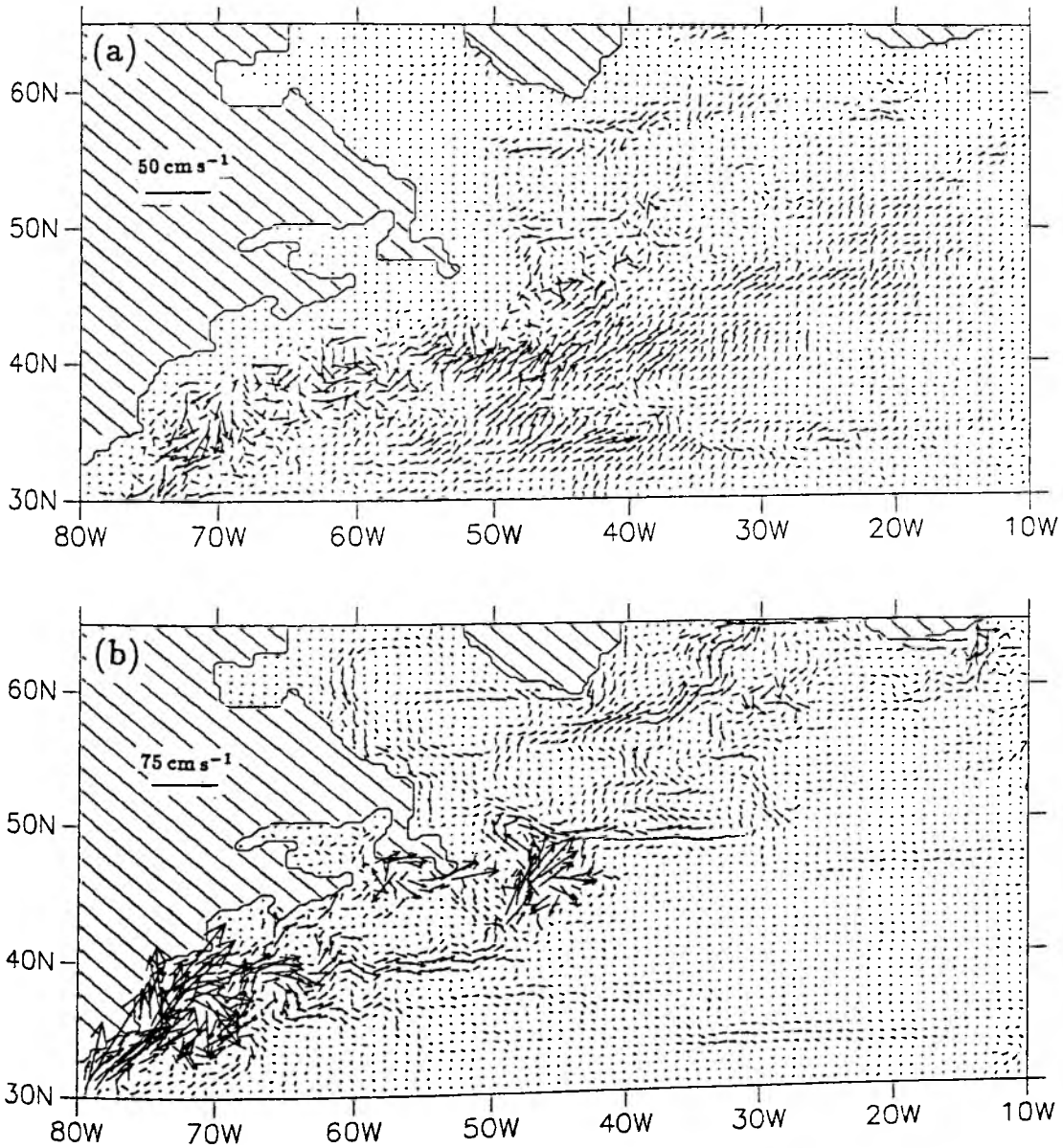


Figure 7.13: Annual mean surface currents (50 m). (a) Difference field, re-initialization experiment minus model run with no data assimilated. (b) Corresponding mean currents of the CME model with no data assimilated. Note the different vector scales.

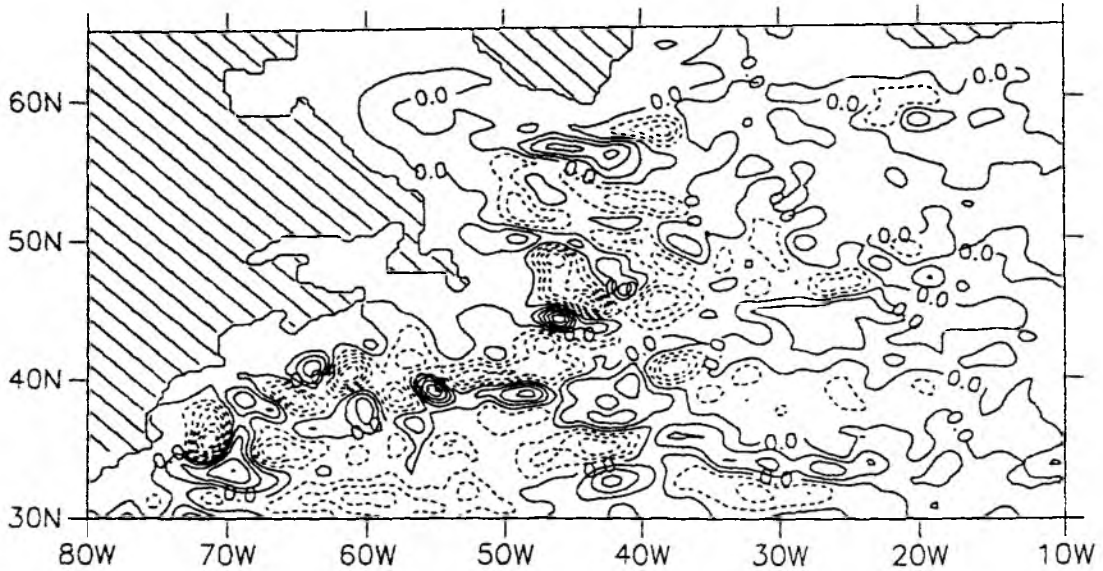


Figure 7.14: *Change in annually averaged SSH between re-initialization experiment and reference run with no data assimilated. The contour interval is 2 cm. Positive values indicate an increase in the mean surface elevation for the assimilation experiment.*

the plots of the actual mean SSH by eye. (The climatological mean surface height of the original CME model was shown in Figure 3.1.) Nevertheless, these changes are consistent with the changes in the mean surface currents investigated above, and they typically point into a direction that is expected to improve the model's climatology. In this respect one might wonder whether using the changed mean surface as reference for a subsequent assimilation experiment, and iteratively repeating this procedure (as suggested by BLAYO *et al.* 1994), would finally converge towards a realistic absolute dynamic topography.

Chapter 8

Conclusions



The present study has demonstrated that the assimilation of satellite altimeter data into a basin-scale, eddy-resolving primitive equation model is not only feasible, but, when applied correctly, can indeed provide new information about the state of today's ocean. Because of the immense computational requirements of so-called optimal methods like the Kalman filter or the adjoint method, one is obviously restricted to adopt much simpler assimilation methods. To still utilize the information in the altimeter data properly, some direct extrapolation of the surface observations into the ocean interior is required. Along these lines, two alternative assimilation methods, referred to as extended nudging and re-initialization, respectively, have been developed. They conceptually differ from each other especially with respect to their vertical extrapolation schemes: The extended nudging approach purely relies on the model statistics to vertically relate the observed surface quantities to the complete set of prognostic model variables. Alternatively, the re-initialization method emphasizes dynamical relationships and kinematic aspects in order to minimize the use of the rather uncertain model statistics.

While both assimilation methods were found to perform similarly well in the context of identical twin experiments, considerable differences emerged when real data, provided by the U.S. Navy's altimetric satellite GEOSAT, were assimilated into the CME model of the North Atlantic Ocean. In particular the use of statistical correlations between sea surface height and temperature/salinity of the numerical model, not being consistent with the statistics of the real ocean as sampled by the GEOSAT altimeter, turned out to be a serious deficiency of the extended nudging approach. Besides creating unrealistic water masses, it considerably interfered with physical pro-

cesses, e.g. by disrupting the seasonal cycle. To summarize these results, it can be said that one should be extremely reluctant to tie uncertain model statics into assimilation procedures, as long as numerical models of the ocean circulation still exhibit considerable systematic deficiencies.

The re-initialization method, on the other hand, was specially constructed to constrain any changes in the model's hydrographic state in a physically sensible way. It explicitly conserves temperature and salinity on isopycnals at individual assimilation steps, thereby avoiding unnecessary distortions of the modelled water mass properties. When judged by the often used statistical measure of explained variance, the performance of this second method actually was slightly worse than that of the extended nudging scheme. However, this only indicates that in order to fully assess the potential role of data assimilation in oceanography, one certainly has to investigate the effect of the assimilation procedure on physically more meaningful quantities, i.e., temperature, salinity and current velocities. Indeed, it turned out that the re-initialization method was able to produce a very reasonable hydrographic picture of the ocean – quite in contrast to the extended nudging scheme that introduced completely unrealistic water masses. The high degree of realism reached by the re-initialization method was further illustrated by the quite satisfactory reproduction of a cold core ring, that agreed very well with independent hydrographic observations reported by KRAUSS *et al.* (1990). Regarding that only sea surface height measurements were used for the assimilation, this is a remarkable success which underlines the additional gain of information resulting from the combination of data with dynamical models.

Besides having the sounder physical basis and producing the more realistic picture of the actual state of the ocean, the re-initialization method has an additional conceptual advantage over the statistical vertical extrapolation scheme of the extended nudging approach, that assumes correct and in particular unbiased model statistics: It does allow for changes in the model climatology. Although a single one-year experiment is too short to fully reveal the effect of this essentially advective phenomenon, the observed changes in the mean fields are quite encouraging: Mean currents are strengthened in the North Atlantic Current, which in the assimilation experiment with data from the year 1987 crosses the Mid-Atlantic Ridge in form of two current bands. On the other hand, the spurious northward current in the Irminger Sea is efficiently reduced by the assimilation of altimeter data. Even the mean surface height field, that serves as reference for the GEOSAT measured SSH anomalies and therefore tends to obstruct large deviations from climatology, seems to improve. Iteratively repeating the assimilation experiment, referencing the GEOSAT data to the newly computed mean

sea surface, may finally lead to a much better surface height climatology, associated with a more realistic representation of major current systems.

Of further interest for future applications is also the property of the re-initialization scheme to conserve temperature and salinity on isopycnals. By improving the model's capability to accurately reproduce location and structure of cyclonic anomalies and frontal systems, the assimilation of altimeter data can inherently control the temporal and spatial setting of water mass formation associated with these dynamical features. With temperature and salinity of the newly formed water essentially being fixed by the atmospheric conditions in the formation area, these properties will not be changed by subsequent re-initialization steps. Combining the assimilation of the expected large quantities of forthcoming accurate altimeter data with more reliable surface boundary conditions may in this way become a valuable tool for studying mesoscale and inter-annual variability in the ocean. It is especially the concept of conserving water mass properties that, I believe, opens a promising way for the assimilation of surface restricted satellite data to eventually contribute to a better understanding of the physical mechanisms driving the ocean.

Appendix A

Observability of Linear Systems

A short excursion to the theory of linear systems is used to introduce the concept of observability. Subsequently, a simple criterion will be derived that allows to test whether an observation campaign can potentially lead to a complete description of the system under investigation.

We begin by considering the linear system

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) \quad , \quad (\text{A.1})$$

what for example could be any linear ocean model. The constant $n \times n$ system matrix \mathbf{A} generates a time series of states $\mathbf{x}(t)$ for any given initial state $\mathbf{x}(0)$. In general, the variables of the state vector \mathbf{x} (e.g., velocity components, temperature or salinity values) are not completely independent from each other, but are interconnected by the system matrix \mathbf{A} . In this way the evolution of the state variables is restricted to the state space Ω , that for linear systems is a linear vector space of dimension

$$\dim \Omega = \text{rank } \mathbf{A} \quad . \quad (\text{A.2})$$

The dynamical system (A.1) is homogeneous. No external forcing terms that would destroy this homogeneity have been taken into account. Indeed, for a linear model the superposition principle restricts the action of such forcing terms to establish a particular solution. In this respect, data assimilation is just aimed at determining the one homogeneous solution that corresponds best with the observations.

Let the observation vector $\mathbf{y}^{obs}(t)$ represent the data obtained from all measurements taken at time t . To see how well the model fits the observations, one first has to map the model state $\mathbf{x}(t)$ onto the observed variables. If this mapping is linear, one may construct the equivalent observation matrix \mathbf{E} and write for the observation vector estimated by the model

$$\mathbf{y}^{est}(t) = \mathbf{E}\mathbf{x}(t) \quad , \quad (\text{A.3})$$

where it has been assumed that the observation system and hence \mathbf{E} are constant in time. For a finite time series of N observations $\mathbf{y}^{obs}(0), \mathbf{y}^{obs}(1), \dots, \mathbf{y}^{obs}(N-1)$ one can then define an extended observation matrix $\hat{\mathbf{E}}$ by

$$\hat{\mathbf{y}}_N^{est} \equiv \begin{pmatrix} \mathbf{y}^{est}(0) \\ \mathbf{y}^{est}(1) \\ \mathbf{y}^{est}(2) \\ \vdots \\ \mathbf{y}^{est}(N-1) \end{pmatrix} = \begin{pmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{A} \\ \mathbf{E}\mathbf{A}^2 \\ \vdots \\ \mathbf{E}\mathbf{A}^{N-1} \end{pmatrix} \mathbf{x}(0) \equiv \hat{\mathbf{E}}_N \mathbf{x}(0) . \quad (\text{A.4})$$

The dynamical system (A.1) is called observable if there is a minimum number, N_0 , of observations such that for any estimated observation time series $\hat{\mathbf{y}}_N^{est}$, with $N \geq N_0$, there exists a unique initial state $\mathbf{x}(0)$, or

$$\forall \mathbf{x}(0), \mathbf{x}'(0) \in \Omega, N \geq N_0 : \hat{\mathbf{E}}_N \mathbf{x}(0) = \hat{\mathbf{E}}_N \mathbf{x}'(0) \implies \mathbf{x}(0) = \mathbf{x}'(0) . \quad (\text{A.5})$$

That is, the system is observable if there is only one solution of the model that is compatible with the (long enough) sequence of estimated observations. The above condition is equivalent to the extended observation matrix $\hat{\mathbf{E}}_N$ having full rank,

$$\text{rank } \hat{\mathbf{E}}_N \equiv \dim \hat{\mathbf{E}}_N(\Omega) = \dim \Omega = \text{rank } \mathbf{A} . \quad (\text{A.6})$$

It is in particular evident that even if a single observation can not uniquely define a corresponding state of the model, this is not necessarily true for a time series of these very same observations.

The rank condition (A.6) provides a general test for observability. However, computing the rank of the extended observation matrix $\hat{\mathbf{E}}_N$ becomes more and more laborious as the length of the time series increases. After all, it is possible to find a maximum number N_{max} of observations beyond which the rank of $\hat{\mathbf{E}}_N$ does not change (CHUI and CHEN, 1989): Let r be the dimension of the state space Ω . By applying the Cayley-Hamilton theorem we can expand \mathbf{A}^r as a linear combination of \mathbf{A}^l , $l = 0, \dots, r-1$ as

$$\mathbf{A}^r = c_0 \mathbf{1} + c_1 \mathbf{A} + c_2 \mathbf{A}^2 + \dots + c_{r-1} \mathbf{A}^{r-1} , \quad (\text{A.7})$$

and we can conclude that $N_{max} = r$. This implies that the system is observable by observations represented by \mathbf{E} if and only if

$$\text{rank } \hat{\mathbf{E}}_r = \text{rank} \begin{pmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{A} \\ \mathbf{E}\mathbf{A}^2 \\ \vdots \\ \mathbf{E}\mathbf{A}^{r-1} \end{pmatrix} = r = \text{rank } \mathbf{A} . \quad (\text{A.8})$$

Even when considering very coarse resolution models of the ocean with a state dimension of $r = O(10^3)$ and $m = O(10^2)$ grid points where sea surface height data are available, the task of computing the rank of the $rm \times r$ matrix $\hat{\mathbf{E}}_r$ will be rather forbidding. An escape route can be found by a theorem derived by HAUTUS (1969) stating the following equivalence:

$$\text{rank } \hat{\mathbf{E}}_r = r \iff \forall \lambda \neq 0, \mathbf{x} : \mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \mathbf{E}\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0} \quad (\text{A.9})$$

Thus, the projection of any (right) eigenvector of the system matrix \mathbf{A} onto the observation space has to be nonzero, except for the nullvector. If all the nonzero eigenvalues are pairwise distinct, the linearity of the mapping induced by \mathbf{E} allows us to restrict this test to one arbitrarily chosen basis set of eigenvectors of \mathbf{A} . However, in the case of multiple eigenvalues one has to make sure that no non-trivial linear combination of states with the same eigenvalue can be in the nullspace of \mathbf{E} . In other words, the dimension of each of the eigenspaces of \mathbf{A} must not be changed by applying the mapping induced by \mathbf{E} . As was pointed out by MILLER (1989) this is because all states belonging to the same eigenspace obey the same dynamics. Hence, these states will never be distinguishable unless they can be distinguished at a single observation time. To illustrate this subtlety, which can easily be ignored (e.g., FUKUMORI *et al.*, 1993), consider any system that admits two waves of different wavevectors and identical frequencies. For observations taken at a single location one can then construct a superposition of these two waves which will result in a persistent zero signal at the observation point, although each wave taken separately would clearly produce a sequence of nonzero observations.

In practise the size of the matrix \mathbf{E} for a single observation will be much smaller than that of the extended observation matrix $\hat{\mathbf{E}}_r$. Because of the resulting considerable reduction of computational cost, HAUTUS' theorem provides an important tool that actually prepares us to test common linear ocean circulation models for their observability by means of satellite altimetry.

While the above considerations are most useful when dealing with time-discrete systems (as all numerical ocean circulation models are), we note that an observability criterion very similar to (A.8) holds for time-continuous linear systems (analytical models of the ocean) of the general form

$$\dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t) \quad (\text{A.10})$$

which is defined to be observable if for each continuous series of observations $\mathbf{y}(t) = \mathbf{E}\mathbf{x}(t)$, $t \geq 0$ there is only one initial state $\mathbf{x}(0)$. Assuming the $n \times n$ matrix \mathbf{A} to be

nondefective, i.e., diagonalizable but not necessarily by unitary transformations, the observability criterion – now including the stationary eigenspace with zero eigenvalues – is (e.g., HAUTUS, 1969)

$$\text{rank } \hat{\mathbf{E}}_n = \text{rank} \begin{pmatrix} \mathbf{E} \\ \mathbf{E} \mathbf{A} \\ \mathbf{E} \mathbf{A}^2 \\ \vdots \\ \mathbf{E} \mathbf{A}^{n-1} \end{pmatrix} = n . \quad (\text{A.11})$$

The corresponding version of theorem (A.9) becomes

$$\text{rank } \hat{\mathbf{E}}_n = n \iff \forall \lambda, \mathbf{x} : \mathbf{A} \mathbf{x} = \lambda \mathbf{x}, \mathbf{E} \mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0} \quad (\text{A.12})$$

which now includes the zero eigenvalues describing the stationary modes of the system (A.10).

A.1 Application to Linear Models of the Ocean

To give some instructive examples, the observability condition developed in the previous section will now be applied to linear models of the ocean. Throughout this section we again will assume that perfect and complete measurements of sea surface height are available.

We begin with one of the simplest ocean models, namely the linearized barotropic shallow-water equations for an ocean of constant depth H on the f -plane. The prognostic equations for the surface elevation η and the east- and northward velocity components u and v are

$$\frac{\partial \eta}{\partial t} + H \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0 \quad (\text{A.13})$$

$$\frac{\partial u}{\partial t} - f v = -g \frac{\partial \eta}{\partial x} \quad (\text{A.14})$$

$$\frac{\partial v}{\partial t} + f u = -g \frac{\partial \eta}{\partial y} \quad (\text{A.15})$$

where g is the acceleration due to gravity. The above system can be further simplified by introducing the horizontal divergence, D , and the vertical component of relative vorticity, ξ ,

$$D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} , \quad \xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} . \quad (\text{A.16})$$

Employing a spatial Fourier transform one obtains for each wavevector \mathbf{k}

$$\frac{d}{dt} \begin{pmatrix} \eta \\ D \\ \xi \end{pmatrix} = \begin{pmatrix} 0 & -H & 0 \\ g\mathbf{k}^2 & 0 & f \\ 0 & -f & 0 \end{pmatrix} \begin{pmatrix} \eta \\ D \\ \xi \end{pmatrix}, \quad (\text{A.17})$$

where η , D and ξ are now components of the Fourier mode under consideration and depend on time only. The system matrix is nondefective and its three distinct eigenvalues are $\lambda_1 = 0$ and $\lambda_{2,3} = \pm i\omega$, $\omega = \sqrt{f^2 + g\mathbf{k}^2 H}$ referring to the stationary geostrophic mode and the pair of inertia-gravity waves, respectively. The three corresponding one-dimensional eigenspaces can be represented by

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -\frac{g\mathbf{k}^2}{f} \end{pmatrix} \eta_0, \quad \mathbf{v}_{2,3} = \begin{pmatrix} 1 \\ \mp i\frac{\omega}{H} \\ -\frac{f}{H} \end{pmatrix} \eta_0. \quad (\text{A.18})$$

For $\eta_0 \neq 0$, each of these eigenvectors has a nonzero projection on the corresponding Fourier components of the sea surface elevation, and by theorem (A.12) we can immediately conclude that the barotropic shallow-water system (A.17) is observable by altimetric measurements. For completeness we note that the introduction of divergence and vorticity, (A.16), essentially filters out pure inertia waves. Hence, our finding of complete observability of the system (A.17) is not inconsistent with the fact that pure inertia waves, which produce no surface elevation at all, can not be observed by altimetry.

We continue with a slightly more realistic ocean model by simply adding a second layer to the barotropic system (A.17). The resulting two-layer linearized shallow-water ocean over flat topography on the f -plane can simulate baroclinic flow and is governed by

$$\frac{d}{dt} \begin{pmatrix} \eta \\ D_1 \\ \xi_1 \\ h \\ D_2 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 0 & -H_1 & 0 & 0 & -H_2 & 0 \\ g\mathbf{k}^2 & 0 & f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -H_2 & 0 \\ \alpha g\mathbf{k}^2 & 0 & 0 & \beta g\mathbf{k}^2 & 0 & f \\ 0 & 0 & 0 & 0 & -f & 0 \end{pmatrix} \begin{pmatrix} \eta \\ D_1 \\ \xi_1 \\ h \\ D_2 \\ \xi_2 \end{pmatrix}, \quad (\text{A.19})$$

where $\alpha = \rho_1/\rho_2$, $\beta = 1 - \alpha$, and $i = 1, 2$ refers to the layer in question, H_i is its mean thickness, D_i and ξ_i denote the horizontal divergence and vertical component of relative vorticity defined analogously to (A.16), and h is the vertical displacement of the interface separating the two fluids of densities ρ_i . In addition to the two nonvanishing complex conjugate eigenvalues pairs corresponding to the two possible frequencies (for internal and external inertia-gravity waves, respectively), there are now two zero eigenvalues, $\lambda_1 = \lambda_2 = 0$. That is, the eigenspace of the stationary geostrophic solutions is

two-dimensional. Hence, for measurements of the surface elevation η only, as described by the observation matrix $\mathbf{E} = (1 \ 0 \ 0 \ 0 \ 0 \ 0)$, it is possible to construct eigenvectors belonging to this geostrophic subspace that cannot be observed. This is illustrated by the particular choice of basisvectors spanning the geostrophic eigenspace,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -\frac{g\mathbf{k}^2}{f} \\ 1 \\ 0 \\ -\frac{g\mathbf{k}^2}{f} \end{pmatrix} \eta_0, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ -\frac{\rho_2 - \rho_1}{\rho_2} \frac{g\mathbf{k}^2}{f} \end{pmatrix} \eta_0. \quad (\text{A.20})$$

For the second eigenvector, $\mathbf{E}\mathbf{v}_2 = 0$ reflects the fact that geostrophic motions restricted to the second layer do not produce any signal at the surface. Because on the f -plane any geostrophic motion is stationary, a time series of altimeter measurements will not be able to discriminate baroclinic and barotropic geostrophic modes. Only the superposition of these modes can be fixed by observing the geostrophic surface currents.

Combining our results for both the one- and the two-layer model, we conclude that, on the f -plane, the altimeter reveals no information about geostrophic currents except for the very top layer of the model. For the class of models considered above, no however ingenious assimilation scheme can extract any information about the deep geostrophic flow.

This at present rather disappointing result is caused by the fact that all geostrophic motions are constant in time, what, in turn is a consequence of the geostrophic degeneracy. Thus, there is some hope that the range of observable modes can be increased by taking into account *quasi*-geostrophic dynamics. This can, for example, be done by applying the mid-latitude β -plane approximation, i.e., letting the Coriolis parameter vary linearly with latitude: $f = f_0 + \beta y$. Then the quasi-geostrophic flow of the barotropic shallow-water system (A.13-A.15) is governed by the prognostic equation

$$\frac{\partial}{\partial t} \left(\nabla^2 \eta - \frac{f_0^2}{gH} \eta \right) + \beta \frac{\partial \eta}{\partial x} = 0 \quad (\text{A.21})$$

for the surface elevation η . Taking the Fourier transform results in $\dot{\eta} = -i\omega\eta$, with ω given by the dispersion relation for barotropic Rossby waves, $\omega = -\beta k / (k^2 + l^2 + f^2/(gH))$. For these waves the observability criterion (A.11) is obviously satisfied. GASPARD and WUNSCH (1989) have used this simple model to assimilate six months of GEOSAT data of the North Atlantic. Thereby, they showed that about 10% of the observed variability could be explained by only five barotropic Rossby waves. However,

this small percentage clearly indicates that in order to extract more information from the data one has to use more realistic models. Therefore, the next step will be to consider baroclinic flow on the β -plane, which, on the f -plane was found to be able to sustain non-observable modes.

In the two-layer case the quasi-geostrophic modes of the linear shallow-water system on the mid-latitude β -plane are governed by

$$\frac{\partial}{\partial t} \left(\nabla^2 \eta - \frac{f_0^2}{gH_1} (\eta - h) \right) + \beta \frac{\partial \eta}{\partial x} = 0 \quad (\text{A.22})$$

$$\frac{\partial}{\partial t} \left(\frac{\rho_1}{\rho_2} \nabla^2 \eta + \frac{\rho_2 - \rho_1}{\rho_2} \nabla^2 h - \frac{f_0^2}{gH_2} h \right) + \beta \left(\frac{\rho_1}{\rho_2} \frac{\partial \eta}{\partial x} + \frac{\rho_2 - \rho_1}{\rho_2} \frac{\partial h}{\partial x} \right) = 0 \quad (\text{A.23})$$

It can be readily seen that no initial state $(\eta, h) \neq (0, 0)$ can produce a continuous sequence of $\eta = 0$ observations. By the action of the stretching term, i.e., the second term in equation (A.22) the vertical motion of the interface acts as a “wave-maker” (BERRY and MARSHALL, 1989) and information propagates vertically. From the theory of linear Rossby waves (e.g., LEBLOND and MYSAK, 1978) one can derive that, for any horizontal wavevector, the frequencies of the corresponding vertical modes are pairwise distinct. We can then deduce that complete observations of the surface elevation will in principle be capable to fully constrain mid-latitude quasi-geostrophic layer models, retrospectively justifying the use of such models for assimilating altimeter data. Along the same line one can for example also explain the fact that meddies, which essentially are subsurface phenomena, can be observed by satellite altimetry (STAMMER *et al.*, 1991).

All the models considered so far have been layer models with both the number of layers and also the densities of the individual fluids prescribed. In order to model isopycnal surfacing and the associated water mass formation, important for a complete description of the interior ocean dynamics, one has to move to more realistic models with active thermodynamics. In a recent paper FUKUMORI *et al.* (1993) have investigated the observability problem for a linearized version of SPEM (Semi-spectral Primitive Equation Model) (HAIDVOGEL *et al.*, 1991), treating density as a prognostic variable. They computed a basis set of eigenvectors of a spherical coarse-grid model of the North Atlantic Ocean and showed that each of these eigenvectors had a non-zero projection on the surface pressure. However, their immediate conclusion that this result in combination with HAUTUS’ (1969) theorem (A.9) should prove observability of SPEM is valid only if there are no multi-dimensional eigenspaces. Since degenerate eigenvalues always indicate some symmetry property of the system, that is likely to

vanish as the complexity of the system increases, we shall just cautiously suppose that both the currents and the density field can be observed by satellite altimetry. Additional information should be used to determine temperature and salinity, since changes in these water mass properties may compensate each other with respect to changes in density and hence are not necessarily dynamically active.

Bibliography

- Arhan, M., and A. Colin de Verdière, 1985: Dynamics of eddy motions in the eastern North Atlantic. *J. Phys. Oceanogr.*, **15**, 153-170.
- Anderson, D. L. T., and J. Willebrand (ed.), 1989: Oceanic circulation models: Combining data and dynamics. Kluwer Academic Publ., Amsterdam, 605 pp.
- Anthes, R. A., 1974: Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, **31**, 702-718.
- Beckmann, A., C. W. Böning, C. Köberle, and J. Willebrand, 1994a: Effects of increased horizontal resolution in a simulation of the North Atlantic Ocean. *J. Phys. Oceanogr.*, in press.
- Beckmann, A., C. W. Böning, B. Brügge, and D. Stammer, 1994b: Generation and role of eddy variability in the central North Atlantic Ocean. *J. Geophys. Res.*, submitted.
- Berry, P., and J. Marshall, 1989: Ocean modelling studies in support of altimetry. *Dyn. Atm. Oc.*, **13**, 269-300.
- Blayo, E., J. Verron, and J. M. Molines, 1994: Real-time assimilation of Topex/Poseidon altimeter data into a circulation model of the North Atlantic. *J. Geophys. Res.*, submitted.
- Böning, C. W., 1989: Influences of rough bottom topography on flow kinematics in an eddy-resolving circulation model. *J. Phys. Oceanogr.*, **19**, 77-97.
- Böning, C. W., R. Döscher, and R. Budich, 1991: Seasonal transport variation in the western subtropical North Atlantic: Experiments with an eddy-resolving model. *J. Phys. Oceanogr.*, **21**, 1271-1289.
- Böning, C. W., and R. Budich, 1992: Eddy dynamics in a primitive equation model: Sensitivity to horizontal resolution and friction. *J. Phys. Oceanogr.*, **22**, 361-381.
- Böning, C. W., 1992: Transportprozesse im subtropischen Nordatlantik: Untersuchungen mit wirbelauflösenden Modellen der windgetriebenen und thermohalinen Zirkulation, Habilitationsschrift, Math.-Nat. Fakultät der Christian-Albrechts-Universität Kiel.
- Böning, C. W., and P. Herrmann, 1994: On the annual cycle of poleward heat transport in the ocean: Results from high-resolution modelling of the North and equatorial Atlantic. *J. Phys. Oceanogr.*, **24**, 91-107.

- Bretherton, F., R. Davies, and C. Fandry, 1976: A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep Sea Res.*, **23**, 559-582.
- Bryan, F. O., and W. R. Holland, 1989: A high-resolution simulation of the wind- and thermohaline-driven circulation in the North Atlantic Ocean. In: *Parameterization of small-scale processes*. Proceedings 'Aha huli'ko'a, Hawaiian Winter Workshop, University of Hawaii, 99-115.
- Bryan, K., 1969: A numerical method for the study of the circulation of the World Ocean. *J. Comput. Phys.*, **4**, 347-376.
- Bryan, K., and M. D. Cox, 1972: An approximate equation of state for numerical models of the ocean circulation. *J. Phys. Oceanogr.*, **2**, 510-514.
- Camp, N. T., and R. L. Elsberry, 1978: Oceanic thermal response to strong atmospheric forcing, 2, The role of one-dimensional processes. *J. Phys. Oceanogr.*, **8**, 215-224.
- Chelton, D. B., M. G. Schlax, D. L. Witter, and J. G. Richman, 1990: Geosat altimeter observations of the surface circulation of the Southern Ocean. *J. Geophys. Res.*, **95**, C10, 17877-17903.
- Cheney, R. E., J. G. Marsh, and B. D. Beckley, 1983: Global mesoscale variability from collinear tracks of Seasat altimeter data. *J. Geophys. Res.*, **88**, C7, 4343-4354.
- Cheney, R. E., B. C. Douglas, E. Khedouri, and C. Szszechowski, 1984: Mesoscale variability between New York and Bermuda from repeated XBT sections and satellite altimetry. *J. Phys. Oceanogr.*, **15**, 252-257.
- Cheney, R. E., B. C. Douglas, R. W. Agreen, L. Miller, D. L. Porter, and N. S. Doyle, 1987: GEOSAT altimeter geophysical data record user handbook. *NOAA Tech. Memo. NOS NGS-46*, Rockville, MD, 32 pp.
- Chui, C. K., and G. Chen, 1989: *Linear Systems and Optimal Control*, Springer-Verlag, 155 pp.
- Cohn, S. E., and D. P. Dee, 1988: Observability of discretized partial differential equations. *SIAM J. Numer. Anal.*, **25**, 586-617.
- Colin de Verdière, A., H. Mercier, and M. Arhan, 1989: Mesoscale variability transition from the western to the eastern Atlantic along 48°N. *J. Phys. Oceanogr.*, **19**, 1149-1170.
- Cox, M. D., 1984: A primitive equation, 3-dimensional model of the ocean. *GFDL Ocean Group Tech. Rep. No. 1*, GFDL/Princeton University.
- Cox, M. D., 1985: An eddy-resolving numerical model of the ventilated thermocline. *J. Phys. Oceanogr.*, **15**, 1312-1324.
- Cox, M. D., 1987: An eddy-resolving numerical model of the ventilated thermocline: Time dependence. *J. Phys. Oceanogr.*, **17**, 1044-1056.
- Dantzler, H. L., 1977: Potential energy maxima in the tropical and subtropical North Atlantic. *J. Phys. Oceanogr.*, **7**, 512-519.

- De Mey, P., and A. R. Robinson, 1987: Assimilation of altimeter data eddy fields in a limited-area quasi-geostrophic model. *J. Phys. Oceanogr.*, **17**, 2280-2293.
- De Mey, P., and Y. Ménard, 1989: Synoptic analysis and dynamical adjustment of GEOS 3 and Seasat altimeter eddy fields in the northwest Atlantic. *J. Geophys. Res.*, **94**, C5, 6221-6231.
- Didden, N., and F. Schott, 1992: Seasonal variations in the western tropical Atlantic: Surface circulation from Geosat altimetry and WOCE model results. *J. Geophys. Res.*, **97**, No. C3, 3529-3541.
- Dombrowsky, E., and P. De Mey, 1992: Continuous assimilation in an open domain of the northeast Atlantic, 1. Methodology and application to AthenA-88. *J. Geophys. Res.*, **97**, No. C6, 9719-9731.
- Draper, N. R., and H. Smith, 1981: Applied Regression Analysis. Sec. Ed., John Wiley & Sons, 709 pp.
- Emery, W. J., 1983: Global summary: Review of eddy phenomena as expressed in temperature measurements. In: A. R. Robinson, ed., *Eddies in Marine Science*, Springer-Verlag, Berlin, 354-375.
- Fu, L.-L., J. Vazquez, and C. Perigaud, 1991: Fitting dynamic models to the Geosat sea level observations in the tropical Pacific Ocean. Part I: A free wave model, *J. Phys. Oceanogr.*, **21**, 798-809.
- Fukumori, I., B. Benveniste, C. Wunsch, and D.B. Haidvogel, 1993: Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother. *J. Phys. Oceanogr.*, **23**, 2162-2181.
- Gaspar, P., and C. Wunsch, 1989: Estimates from altimeter data of barotropic Rossby waves in the northwestern Atlantic Ocean. *J. Phys. Oceanogr.*, **19**, 1821-1844.
- Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Application of estimation theory to numerical weather prediction. In: L. Bengtson, M. Ghil, E. Källén, ed., *Dynamic Meteorology: Data Assimilation Methods*, Springer-Verlag, 139-224.
- Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics*, **33**, 141-265.
- Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press, London, 662 pp.
- Haidvogel, D. B., J. Wilkin, and R. Young, 1991: A semi-spectral primitive equation ocean circulation model using vertical sigma and orthogonal curvilinear coordinates. *J. Comput. Phys.*, **94**, 151-185.
- Haines, K., 1991: A direct method for assimilating sea surface height data into ocean models with adjustment to the deep circulation. *J. Phys. Oceanogr.*, **21**, 843-868.
- Haines, K., P. Malanotte-Rizzoli, R. E. Young, and W. R. Holland, 1993: A comparison of two methods for the assimilation of altimeter data into a shallow water model. *Dyn. Atmos. Oceans*, **17**, 89-133.

- Han, Y.-J., 1984: A numerical world ocean circulation model. Part II: A baroclinic experiment. *Dyn. Atmos. Oceans*, **8**, 141-172.
- Hautus, M. L. J., 1969: Controllability and observability conditions of linear autonomous systems. *Proc. of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A*, 443-448.
- Hellerman, S., and M. Rosenstein, 1983: Normal monthly wind stress over the World Ocean with error estimates. *J. Phys. Oceanogr.*, **13**, 1093-1104.
- Heywood, K. J., E. L. McDonagh, and M. A. White, 1994: Eddy kinetic energy of the North Atlantic subpolar gyre from satellite altimetry. *J. Geophys. Res.*, submitted.
- Hoke, J. E., and R. A. Anthes, 1976: The initialization of numerical models by a dynamic initialization technique. *Mon. Wea. Rev.*, **104**, 1551-1556.
- Holland, W. R., 1989: Altimeter data assimilation into ocean circulation models - some preliminary results. In: *D. L. T. Anderson and J. Willebrand, ed., Oceanic Circulation Models: Combining Data and Dynamics*, Kluwer Academic Publ., 203-231.
- Holland, W. R., and P. Malanotte-Rizzoli, 1989: Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies. *J. Phys. Oceanogr.*, **19**, 1507-1534.
- Isemer, H.-J., and L. Hasse, 1987: *The Bunker Climate Atlas of the North Atlantic Ocean, Vol. 2: Air-Sea Interactions*. Springer-Verlag, 252 pp.
- Kindle, J. C., 1986: Sampling strategies and model assimilation of altimetric data for ocean monitoring and prediction. *J. Geophys. Res.*, **91**, C2, 2418-2432.
- Krauß, W., and R. H. Käse, 1984: Mean circulation and eddy kinetic energy in the eastern North Atlantic. *J. Geophys. Res.*, **89**, C3, 3407-3415.
- Krauß, W., R. H. Käse, and H.-H. Hinrichsen, 1990: The branching of the Gulf Stream southeast of the Grand Banks. *J. Geophys. Res.*, **95**, C8, 13089-13103.
- Le Blond, P. H., and L. A. Mysak, 1978: *Waves in the Ocean*. Elsevier Scientific Publishing Co., Amsterdam, 602 pp.
- Le Dimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97-110.
- Le Traon, P. Y., M. C. Rouquet, and C. Boissier, 1990: Spatial scales of mesoscale variability in the North Atlantic as deduced from Geosat data. *J. Geophys. Res.*, **95**, C11, 20267-20285.
- Le Traon, P. Y., 1991: Time scales of mesoscale variability and their relationship with space scales in the North Atlantic. *J. Mar. Res.*, **49**, 467-492.
- Le Traon, P. Y., 1992: Contribution of satellite altimetry to the observation of oceanic mesoscale variability. *Ocean. Acta*, **15**, 5, 441-457.
- Levitus, S., 1982: Climatological atlas of the world ocean. *NOAA Prof. Pap.* 13, U.S. Govt. Print. Office, Washington, D.C., 173 pp.

- Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309-322.
- Malanotte-Rizzoli, P., R. E. Young, and D. B. Haidvogel, 1989: Initialization and data assimilation experiments with a primitive equation model. *Dyn. Atm. Oc.*, **13**, 349-378.
- Marotzke, J., and C. Wunsch, 1993: Finding the steady state of a general circulation model through data assimilation: Application to the North Atlantic Ocean. *J. Geophys. Res.*, **98**, C11, 20149-20167.
- Mellor, G. L., and T. Ezer, 1991: A Gulf Stream model and an altimetry assimilation scheme. *J. Geophys. Res.*, **96**, C5, 8779-8795.
- Miller, R. N., 1989: Direct assimilation of altimetric differences using the Kalman filter. *Dyn. Atm. Oc.*, **13**, 317-333.
- Müller, P., and J. Willebrand, 1989: Equations for oceanic motions. In: *J. Sündermann (Ed.), Landolt-Börnstein, Group V, Oceanography, Volume 3b*, Springer Verlag, Berlin, 1-14.
- Richardson, P. L., 1983: Eddy kinetic energy in the North Atlantic from surface drifters. *J. Geophys. Res.*, **88**, C7, 4355-4367.
- Robinson, A. R., M. A. Spall, and N. Pinardi, 1988: Gulf Stream simulations and the dynamics of ring and meander processes. *J. Phys. Oceanogr.*, **18**, 1811-1853.
- Sailor, R. V., and A. R. Le Schack, 1987: Preliminary determination of the Geosat radar altimeter noise spectrum. *Johns Hopkins APL Tech. Dig.*, **8(2)**, 182-183.
- Schott, F. A., and C. W. Böning, 1991: Evaluation of the WOCE model in the western equatorial Atlantic: Upper-layer circulation. *J. Geophys. Res.*, **96**, C4, 6993-7004.
- Schröter, J., U. Seiler, and M. Wenzel, 1993: Variational assimilation of Geosat data into an eddy-resolving model of the Gulf Stream extension area. *J. Phys. Oceanogr.*, **23**, 925-953.
- Semtner, A. J., 1986: Finite-difference formulation of a World Ocean model. In *J. J. O'Brien, ed., Advanced Physical Oceanographic Numerical Modelling*, D. Reidel Publishing Co., Dordrecht, 187-202.
- Spall, M. A., 1990: Circulation in the Canari Basin: A model/data analysis. *J. Geophys. Res.*, **95**, C12, 15502-15522.
- Stammer, D., H.-H. Hinrichsen, and R. H. Käse, 1991: Can Meddies be detected by satellite altimetry? *J. Geophys. Res.*, **96**, C4, 7005-7014.
- Stammer, D., 1992: Über die mesoskalige Variabilität im Atlantischen Ozean, Analyse und Assimilation von GEOSAT-Altimeterdaten. *Ber. Inst. f. Meeresk. Kiel*, Nr. 224, 197 pp.

- Stammer, D., and C. W. Böning, 1992: Mesoscale variability in the Atlantic Ocean from GEOSAT Altimetry and WOCE high resolution numerical modeling. *J. Phys. Oceanogr.*, **22**, 732-752.
- Stensrud, D. J., and J. W. Bao, 1992: Behaviors of variational and nudging assimilation techniques with a chaotic low-order model. *Mon. Wea. Rev.*, **120**, 3016-3028.
- Sy, A., U. Schauer, and J. Meincke, 1992: The North Atlantic Current and its associated hydrographic structure above and eastwards of the Mid-Atlantic Ridge. *Deep Sea Res.*, **39**, 825-853.
- Thacker, W. C., 1986: Relationships between statistical and deterministic methods of data assimilation. In: Y. K. Sasaki, ed., *Variational Methods in Geosciences*, Elsevier Scientific Publishing Co., Amsterdam, 173-179.
- Treguier, A.M., 1992: Kinetic energy analysis of an eddy resolving, primitive equation model of the North Atlantic. *J. Geophys. Res.*, **97**, C1, 687-701.
- Verron, J., 1992: Nudging satellite altimeter data in quasi-geostrophic ocean models. *J. Geophys. Res.*, **97**, C5, 7479-7491.
- Verron, J., J. M. Molines, and E. Blayo, 1992: Assimilation of Geosat data into a quasigeostrophic model of the North Atlantic between 20°N and 50°N: preliminary results. *Oceanol. Acta*, **15**, 575-583.
- White, W. B., C.-K. Tai, and W. R. Holland, 1990a: Continuous assimilation of Geosat altmetric sea level observations into a numerical synoptic ocean model of the California Current. *J. Geophys. Res.*, **95**, C3, 3127-3148.
- White, W. B., C.-K. Tai, and W. R. Holland, 1990b: Continuous assimilation of simulated Geosat altmetric sea level into an eddy-resolving numerical ocean model, 1. sea level differences. *J. Geophys. Res.*, **95**, C3, 3219-3234.
- White, W. B., C.-K. Tai, and W. R. Holland, 1990c: Continuous assimilation of simulated Geosat altmetric sea level into an eddy-resolving numerical ocean model, 1. referenced sea level differences. *J. Geophys. Res.*, **95**, C3, 3235-3251.
- Willebrand, J., R. H. Käse, D. Stammer, H.-H. Hinrichsen, and W. Krauß, 1990: Verification of Geosat sea surface topography in the Gulf Stream extension with surface drifting buoys and hydrographic measurements. *J. Geophys. Res.*, **95**, C3, 3007-3014.
- Zou, X., I. M. Navon, and F. X. Le Dimet, 1992: Incomplete observations and control of gravity waves in variational data assimilation. *Tellus*, **44A**, 273-296.

Acknowledgments

I would like to thank my supervisor, Prof. Dr. Jürgen Willebrand, for his guidance and support as well as the independence he confidentially allowed me. Dr. Detlef Stammer kindly provided the preprocessed GEOSAT data, which were mapped by a routine originally developed by Dr. Pierre De Mey. Many thanks are due to Peter Herrmann for his valuable help in running the CME model experiments. Finally it is a pleasure to thank my colleagues at the Institut für Meereskunde in Kiel and also Dr. Keith Haines of Edinburgh for many helpful and encouraging conversations.